

# SECOND-ORDER SUBDIFFERENTIAL CALCULUS WITH APPLICATIONS TO TILT STABILITY IN OPTIMIZATION

B. S. MORDUKHOVICH<sup>1</sup> and R. T. ROCKAFELLAR<sup>2</sup>

**Abstract.** The paper concerns the second-order generalized differentiation theory of variational analysis and new applications of this theory to some problems of constrained optimization in finite-dimensional spaces. The main attention is paid to the so-called (full and partial) second-order subdifferentials of extended-real-valued functions, which are dual-type constructions generated by coderivatives of first-order subdifferential mappings. We develop an extended second-order subdifferential calculus and analyze the basic second-order qualification condition ensuring the fulfillment of the principal second-order chain rule for strongly and fully amenable compositions. The calculus results obtained in this way and computing the second-order subdifferentials for piecewise linear-quadratic functions and their major specifications are applied then to the study of tilt stability of local minimizers for important classes of problems in constrained optimization that include, in particular, problems of nonlinear programming and certain classes of extended nonlinear programs described in composite terms.

**Key words.** variational analysis, constrained optimization, nonlinear and extended nonlinear programming, second-order subdifferentials, calculus rules, qualification conditions, amenable functions, tilt-stable minimizers, strong regularity

**AMS subject classifications.** 49J52, 90C30, 90C31

**Abbreviated title.** Second-order subdifferential calculus

## 1 Introduction

Variational analysis has been recognized as a fruitful area of mathematics, which primarily deals with optimization-related problems while also applying variational principles and techniques (largely based on perturbation and approximation ideas) to a broad spectrum of problems that may not be of a variational nature. We refer the reader to the books by Borwein and Zhu [3], Mordukhovich [29, 30], Rockafellar and Wets [45], and the bibliographies therein for the major results of variational analysis and its numerous applications.

Since nonsmooth functions, sets with nonsmooth boundaries, and set-valued mappings naturally and frequently appear in the framework of variational theory and its applications via using variational principles and techniques (even for problems with smooth initial data), tools of generalized differentiation play a crucial role in many aspects of variational analysis and optimization; see, e.g., the books [3, 7, 10, 17, 29, 30, 45, 46] and their references.

Over the years, the first-order subdifferential theory of variational analysis has been well developed and understood in both finite-dimensional and infinite-dimensional settings; see [3, 29, 45] and the commentaries therein. In contrast, the *second-order* theory still requires a lot of further development and implementation, although many second-order generalized differential constructions have been suggested and successfully applied to various optimization, sensitivity,

---

<sup>1</sup>Department of Mathematics, Wayne State University, Detroit, MI 48202 (boris@math.wayne.edu). Research of this author was partly supported by the National Science Foundation under grant DMS-1007132, by the Australian Research Council under grant DP-12092508, and by the Portuguese Foundation of Science and Technologies under grant MAT/11109.

<sup>2</sup>Department of Mathematics, University of Washington, Seattle, WA 98195 (rtr@math.washington.edu).

and related problems; see, e.g., the books [2, 29, 45] summarizing mainstream developments and trends in the second-order theory and its applications.

As is well known, there are two generally independent approaches to second-order differentiation in the classical analysis. One of them is based on the Taylor expansion while the other defines the second derivative of a function as the derivative of its first-order derivative.

In this paper we develop the latter “derivative-of-derivative” approach to the second-order generalized differentiation of extended-real-valued functions  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := (-\infty, \infty]$  finite at the reference points. The dual-space route in this vein suggested by Mordukhovich [24] is to treat a (first-order) *subdifferential*  $\partial\varphi$  of  $\varphi$  as a set-valued analog of the classical derivative for nonsmooth functions and then to define a *second-order subdifferential*  $\partial^2\varphi$  of  $\varphi$  via a *coderivative* (generalized adjoint derivative operator)  $D^*\partial\varphi$  of the subgradient mapping  $\partial\varphi$ ; see Section 2 for more details. This second-order construction was originally motivated by applications to sensitivity analysis of variational systems [24, 27] inspired by the coderivative characterization of Lipschitzian stability [24, 25], but then the second-order subdifferential and its modification were successfully employed in the study of a broad spectrum of other important issues in variational analysis and its applications; see, e.g., [4, 9, 11, 12, 13, 15, 17, 18, 19, 29, 30, 31, 32, 36, 39, 50, 51] and the references therein. We specifically mention a remarkable result by Poliquin and Rockafellar [39] who established a full characterization of *tilt-stable local minimizers* of functions (a new notion introduced by them motivated, in particular, by the justification of numerical algorithms) as the positive-definiteness of the second-order subdifferential mapping. For  $\mathcal{C}^2$  functions, the latter criterion reduces to the positive-definiteness of the classical Hessian matrix—a well-known sufficient condition for the standard optimality in unconstrained problems, which happens to be necessary and sufficient for tilt-stable local minimizers [39]. We also refer the reader to the recent papers by Chieu et al. [4, 5] providing complete characterizations of *convexity* and *strong convexity* of nonsmooth (in the second order) functions via positive-semidefiniteness and definiteness of their second-order subdifferentials  $\partial^2\varphi$ . Related characterizations of *monotonicity* and *submonotonicity* of continuous mappings can be found in [6].

Needless to say that efficient implementations and potential extensions of the latter result to constrained optimization problems, as well as any other valuable applications of the aforementioned second-order subdifferential construction and its modifications, largely depend on the possibility to develop a fairly rich *second-order subdifferential calculus* and on precisely *calculating* such constructions for attractive classes of nonsmooth functions overwhelmingly encountered in variational analysis and optimization. A certain amount of useful second-order calculus rules were developed in [18, 26, 28, 29, 31, 32, 34]. On the other hand, precisely calculating the second-order subdifferential entirely in terms of the initial data was effected for the following major classes of extended-real-valued functions particularly important in various applications:

- For the indicator functions of *convex polyhedra* and related settings it was initiated by Dontchev and Rockafellar [9] and then developed in [1, 12, 13, 15, 16, 35, 40, 48] for more involved frameworks. The obtained calculations played a crucial role in deriving [9] verifiable characterizations of Robinson’s strong regularity [41] for variational inequalities over (convex) polyhedral sets as well as their specifications for complementarity problems and the associated Karush-Kuhn-Tucker (KKT) conditions for nonlinear programs with  $\mathcal{C}^2$  data. Further results in this vein on Lipschitzian stability of parametric variational systems were given in [1, 12, 35, 40, 49] and other publications in both finite and infinite dimensions. Applications to stationarity conditions for stochastic equilibrium problems with equilibrium constraints in electricity spot market modeling were developed by Henrion and Römisch [15].

- For the so-called *separable piecewise  $\mathcal{C}^2$*  functions it was done by Mordukhovich and Outrata [31]; see also [4] for further developments. it provided the basis for the efficient sensitivity analysis

[31] of mathematical programs with equilibrium constraints (MPECs) including practical ones that arose in applications to certain contact problems of continuum mechanics.

- For indicator functions to *smooth nonpolyhedral inequality systems* it was done by Henrion, Outrata, and Surowiec [13] by employing and developing the transformation formulas from [31]. Then these calculations were applied in [14, 47] to deriving stationarity conditions for equilibrium problems with equilibrium constraints (EPECs) in both deterministic and stochastic frameworks and to EPEC models of oligopolistic competition in electricity spot markets.

- For the special class of functions arising in optimal control of the Moreau *sweeping process* it was done in the paper by Colombo et al. [8]. These calculations played a significant role in deriving constructive optimality conditions for discontinuous differential inclusions generated by the sweeping process with great potentials for further applications.

Now we briefly describe the main *goals and achievements* of this paper. Our primary attention is focused on the following major issues new in second-order variational analysis:

- Developing refined *second-order chain rules* of the equality and inclusion (outer/upper estimate) types for the aforementioned second-order subdifferential and its partial modifications.
- Analyzing the basic *second-order qualification condition* ensuring the fulfillment of the extended second-order chain rules for *strongly amenable* compositions.
- Precise *calculating* second-order subgradients for major classes of *fully amenable* functions.
- Applications of the obtained calculus and computational results to deriving necessary optimality conditions as well as to establishing *complete characterizations* of *tilt-stable minimizers* for broad classes of constrained optimization problems including those in nonlinear programming (NLP) and extended nonlinear programming (ENLP) described via amenable compositions.

The rest of the paper is organized as follows. Section 2 contains basic definitions and brief discussions of the first-order and second-order generalized differential constructions studied and used in the paper. We also review there some preliminary results widely employed in the sequel.

In Section 3 we deal with *second-order chain rules* of the equality and inclusion types for the basic second-order subdifferential and its partial counterparts. The equality-type results are established under the *full rank condition* on the Jacobian matrix of the inner mapping of the composition. Without imposing the latter assumption, we develop a new *quadratic penalty* approach that allows us to derive inclusion-type second-order chain rules for a broad class of strongly amenable compositions valid under certain second-order qualification conditions. The latter chain rules generally provides merely outer estimates of the second-order subgradient sets for compositions: we present an example showing that the chain rule inclusion may be strict even the linear inner mapping and piecewise linear outer functions in fully amenable compositions.

Section 4 is devoted to a detailed analysis of the basic *second-order qualification condition* ensuring the underlying second-order chain rule for strongly amenable compositions. Although the latter condition is automatically fulfilled under the full rank assumption on the Jacobians of inner mappings as well as for  $C^{1,1}$  outer functions in compositions, it seems to be rather restrictive when outer functions are extended-real-valued. In particular, we show that the second-order qualification condition *reduces locally* to the *full rank* requirement on the inner mapping Jacobian matrix if the outer function is either convex piecewise linear, or it belongs to a certain major class of piecewise linear-quadratic functions. The results obtained in this direction are based on precise *calculations* of the second-order subgradient sets of the remarkable classes of fully amenable compositions under consideration.

The concluding Section 5 concerns applications of the chain rules and calculation results developed in the previous sections to the study of *tilt-stable minimizers* for some classes of constrained optimization problems represented in composite formats, which are convenient for developing both theoretical and computational aspects of optimization. Such classes include,

besides standard nonlinear programs (NLP), broader models of the so-called extended nonlinear programming (ENLP). Based on the second-order sum and chain rules with equalities, we derive *complete characterizations* of tilt-stable local minimizers for important problems of constrained optimization. The results obtained show, in particular, that for a general class of NLP problems the well-recognized *strong second-order optimality condition* is *necessary and sufficient* for the tilt-stability of local minimizers, which therefore is equivalent to Robinson's *strong regularity* of the associated variational inequalities in such settings. Furthermore, the calculus rules obtained in this paper for *partial* second-order subdifferentials lead us also to characterizations of *full stability* in optimization (see Remark 5.6), while a detailed elaboration of this approach is a subject of our ongoing research.

Although a number of the results obtained in this paper hold in (or can be naturally extended to) infinite-dimensional spaces, we confine ourselves for definiteness to the finite-dimensional setting. Throughout the paper we use standard notation of variational analysis; cf. [29, 45]. Recall that, given a set-valued mapping  $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , the symbol

$$(1.1) \quad \limsup_{x \rightarrow \bar{x}} F(x) := \left\{ y \in \mathbb{R}^m \mid \begin{array}{l} \exists x_k \rightarrow \bar{x}, \exists y_k \rightarrow y \text{ as } k \rightarrow \infty \\ \text{with } y_k \in F(x_k) \text{ for all } k \in \mathbb{N} := \{1, 2, \dots\} \end{array} \right\}$$

signifies the *Painlevé-Kuratowski outer/upper limit* of  $F$  as  $x \rightarrow \bar{x}$ . Given a set  $\Omega \subset \mathbb{R}^n$  and an extended-real-valued function  $\varphi: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  finite at  $\bar{x}$ , the symbols  $x \xrightarrow{\Omega} \bar{x}$  and  $x \xrightarrow{\varphi} \bar{x}$  stand for  $x \rightarrow \bar{x}$  with  $x \in \Omega$  and for  $x \rightarrow \bar{x}$  with  $\varphi(x) \rightarrow \varphi(\bar{x})$ , respectively.

## 2 Basic Definitions and Preliminaries

In this section we define and briefly discuss the basic generalized differential constructions of our study and review some preliminaries widely used in what follows; see [29, 45] for more details.

Let  $\varphi: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  be an extended-real-valued function finite at  $\bar{x}$ . The *regular subdifferential* (known also as the presubdifferential and as the Fréchet or viscosity subdifferential) of  $\varphi$  at  $\bar{x}$  is

$$(2.1) \quad \widehat{\partial}\varphi(\bar{x}) := \left\{ v \in \mathbb{R}^n \mid \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \right\}.$$

Each  $v \in \widehat{\partial}\varphi(\bar{x})$  is a *regular subgradient* of  $\varphi$  at  $\bar{x}$ . While  $\widehat{\partial}\varphi(\bar{x})$  reduces to a singleton  $\{\nabla\varphi(\bar{x})\}$  if  $\varphi$  is Fréchet differentiable at  $\bar{x}$  with the gradient  $\nabla\varphi(\bar{x})$  and to the classical subdifferential of convex analysis if  $\varphi$  is convex, the set (2.1) may often be empty for nonconvex and nonsmooth functions as, e.g., for  $\varphi(x) = -|x|$  at  $\bar{x} = 0 \in \mathbb{R}$ . Another serious disadvantage of the subdifferential construction (2.1) is the failure of standard calculus rules inevitably required in the theory and applications of variational analysis and optimization. In particular, the inclusion (outer estimate) sum rule  $\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) + \widehat{\partial}\varphi_2(\bar{x})$  does not hold for the simplest nonsmooth functions  $\varphi_1(x) = |x|$  and  $\varphi_2(x) = -|x|$  at  $\bar{x} = 0 \in \mathbb{R}$ .

The picture dramatically changes when we employ a limiting “robust regularization” procedure over the subgradient mapping  $\widehat{\partial}\varphi(\cdot)$  that leads us to the (basic first-order) *subdifferential* of  $\varphi$  at  $\bar{x}$  defined by

$$(2.2) \quad \partial\varphi(\bar{x}) := \limsup_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x)$$

and known also as the general, or limiting, or Mordukhovich subdifferential; it was first introduced in [22] in an equivalent way. Each  $v \in \partial\varphi(\bar{x})$  is called a (basic) *subgradient* of  $\varphi$  at  $\bar{x}$ .

Thus, by taking into account definition (1.1) of Limsup and the notation  $x \xrightarrow{\varphi} \bar{x}$ , we represent the basic subgradients  $v \in \partial\varphi(\bar{x})$  as follows:

there are sequences  $x_k \rightarrow \bar{x}$  with  $\varphi(x_k) \rightarrow \varphi(\bar{x})$  and  $v_k \in \widehat{\partial}\varphi(x_k)$  with  $v_k \rightarrow v$ .

In contrast to (2.1), the subgradient set (2.2) is generally nonconvex (e.g.,  $\partial\varphi(0) = \{-1, 1\}$  for  $\varphi(x) = -|x|$ ) while enjoying comprehensive calculus rules (“full calculus”); this is based on *variational/extremal principles*, which replace separation arguments in the absence of convexity. Moreover, the basic subdifferential (2.2) occurs to be the *smallest* among any axiomatically defined subgradient sets satisfying certain natural requirements; see [33, Theorem 9.7].

In what follows we also need another subdifferential construction effective for non-Lipschitzian extended-real-valued functions. Given  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  finite at  $\bar{x}$ , the *singular/horizontal subdifferential*  $\partial^\infty\varphi(\bar{x})$  of  $\varphi$  at  $\bar{x}$  is defined by

$$(2.3) \quad \partial^\infty\varphi(\bar{x}) := \text{Lim sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda \downarrow 0}} \lambda \widehat{\partial}\varphi(x).$$

If the function  $\varphi$  is lower semicontinuous (l.s.c.) around  $\bar{x}$ , then  $\partial^\infty\varphi(\bar{x}) = \{0\}$  if and only if  $\varphi$  is locally Lipschitzian around this point.

Given further a nonempty subset  $\Omega \subset \mathbb{R}^n$ , consider its indicator function  $\delta(x; \Omega)$  equal to 0 for  $x \in \Omega$  and to  $\infty$  otherwise. For any fixed  $\bar{x} \in \Omega$ , define the *regular normal cone* to  $\Omega$  at  $\bar{x}$  by

$$(2.4) \quad \widehat{N}(\bar{x}; \Omega) := \widehat{\partial}\delta(\bar{x}; \Omega) = \left\{ v \in \mathbb{R}^n \mid \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}$$

and similarly the (basic, limiting) *normal cone* to  $\Omega$  at  $\bar{x}$  by  $N(\bar{x}; \Omega) := \partial\delta(\bar{x}; \Omega)$ . It follows from (2.2) and (2.4) that the normal cone  $N(\bar{x}; \Omega)$  admits the limiting representation

$$(2.5) \quad N(\bar{x}; \Omega) = \text{Lim sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x; \Omega)$$

meaning that the basic normals  $v \in N(\bar{x}; \Omega)$  are those vectors  $v \in \mathbb{R}^n$  for which there are sequences  $x_k \rightarrow \bar{x}$  and  $v_k \rightarrow v$  with  $x_k \in \Omega$  and  $v_k \in \widehat{N}(x_k; \Omega)$ ,  $k \in \mathbb{N}$ . If  $\Omega$  is locally closed around  $\bar{x}$ , (2.5) is equivalent to the original definition by Mordukhovich [22]:

$$N(\bar{x}; \Omega) = \text{Lim sup}_{x \rightarrow \bar{x}} \left[ \text{cone}(x - \Pi(x; \Omega)) \right],$$

where  $\Pi(x; \Omega)$  signifies the Euclidean projector of  $x \in \mathbb{R}^n$  on the set  $\Omega$ , and where “cone” stands for the conic hull of a set.

There is the duality/polarity correspondence

$$(2.6) \quad \widehat{N}(\bar{x}; \Omega) = T(\bar{x}; \Omega)^* := \left\{ v \in \mathbb{R}^n \mid \langle v, w \rangle \leq 0 \text{ for all } w \in T(\bar{x}; \Omega) \right\}$$

between the regular normal cone (2.4) and the *tangent cone* to  $\Omega$  at  $\bar{x} \in \Omega$  defined by

$$(2.7) \quad T(\bar{x}; \Omega) := \left\{ w \in \mathbb{R}^n \mid \exists x_k \in \Omega, \exists \alpha_k \geq 0 \text{ with } \alpha_k(x_k - \bar{x}) \rightarrow w \text{ as } k \rightarrow \infty \right\}$$

and known also as the Bouligand-Severi contingent cone to  $\Omega$  at this point. Note that the basic normal cone (2.5) cannot be tangentially generated in a polar form (2.6) by using some set of tangents, since it is intrinsically nonconvex while the polar  $T^*$  to any set  $T$  is automatically

convex. In what follows we may also use the subindex set notation like  $N_\Omega(\bar{x})$ ,  $T_\Omega(\bar{x})$ , etc. for the constructions involved.

It is worth observing that the *convex closure*

$$(2.8) \quad \overline{N}(\bar{x}; \Omega) := \text{clco } N(\bar{x}; \Omega)$$

of (2.5), known as the Clarke/convexified normal cone to  $\Omega$  at  $\bar{x}$  (see [7]), may dramatically enlarge the set of basic normals (2.5). Indeed, it is proved by Rockafellar [43] that for every vector function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  locally Lipschitzian around  $\bar{x}$  the convexified normal cone (2.8) to the graph of  $f$  at  $(\bar{x}, f(\bar{x}))$  is in fact a *linear subspace* of dimension  $d \geq m$  in  $\mathbb{R}^n \times \mathbb{R}^m$ , where the equality  $d = m$  holds if and only if the function  $f$  is strictly differentiable at  $\bar{x}$  with the derivative (Jacobian matrix) denoted for simplicity by  $\nabla f(\bar{x})$ , i.e.,

$$\lim_{x, u \rightarrow \bar{x}} \frac{f(x) - f(u) - \nabla f(\bar{x})(x - u)}{\|x - u\|} = 0,$$

which is automatic when  $f$  is  $\mathcal{C}^1$  around  $\bar{x}$ . In particular, this implies that  $\overline{N}((\bar{x}, f(\bar{x})); \text{gph } f)$  is the whole space  $\mathbb{R}^n \times \mathbb{R}^m$  whenever  $f$  is nonsmooth around  $\bar{x}$ ,  $n = 1$ , and  $m \geq 1$ . Moreover, the aforementioned results are discovered by Rockafellar [43] not only for graphs of locally Lipschitzian functions, but also for the so-called *Lipschitzian manifolds* (or graphically Lipschitzian sets), which are locally homeomorphic to graphs of Lipschitzian vector functions. The latter class includes graphs of maximal monotone relations and subdifferential mappings for convex, saddle, lower- $\mathcal{C}^2$ , and more general prox-regular functions typically encountered in variational analysis and optimization. In fact such *graphical sets* play a crucial role in the coderivative and second-order subdifferential constructions studied in this paper.

Given a set-valued mapping  $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , define its *coderivative* at  $(\bar{x}, \bar{y}) \in \text{gph } F$  by [23]

$$(2.9) \quad D^*F(\bar{x}, \bar{y})(v) := \left\{ u \in \mathbb{R}^n \mid (u, -v) \in N((\bar{x}, \bar{y}); \text{gph } F) \right\}, \quad v \in \mathbb{R}^m,$$

via the normal cone (2.5) to the graph  $\text{gph } F$ . Clearly the mapping  $D^*F(\bar{x}, \bar{y}): \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  is positive-homogeneous; it reduces to the adjoint derivative

$$(2.10) \quad D^*F(\bar{x})(v) = \{\nabla F(\bar{x})^*v\}, \quad v \in \mathbb{R}^m,$$

where  $*$  stands for the matrix transposition, if  $F$  is single-valued (then we omit  $\bar{y} = F(\bar{x})$  in the coderivative notation) and strictly differentiable at  $\bar{x}$ . Note that the coderivative values in (2.9) are often nonconvex sets due to the nonconvexity of the normal cone on the right-hand side. Furthermore, the latter cone is taken to a graphical set, and thus its convexification in (2.9) may create serious troubles; see above.

The main construction studied in the paper was introduced in [24] as follows.

**Definition 2.1 (second-order subdifferential).** *Let the function  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be finite at  $\bar{x}$ , and let  $\bar{y} \in \partial\varphi(\bar{x})$  be a basic first-order subgradient of  $\varphi$  at  $\bar{x}$ . Then the SECOND-ORDER SUBDIFFERENTIAL of  $\varphi$  at  $\bar{x}$  relative to  $\bar{y}$  is defined by*

$$(2.11) \quad \partial^2\varphi(\bar{x}, \bar{y})(u) := (D^*\partial\varphi)(\bar{x}, \bar{y})(u), \quad u \in \mathbb{R}^n,$$

via the coderivative (2.9) of the first-order subdifferential mapping (2.2).

Observe that if  $\varphi \in \mathcal{C}^2$  around  $\bar{x}$  (in fact, it is merely continuously differentiable around  $\bar{x}$  with the strict differentiable first-order derivative at this point), then

$$\partial^2\varphi(\bar{x})(u) = \{\nabla^2\varphi(\bar{x})u\}, \quad u \in \mathbb{R}^n,$$

where  $\nabla^2\varphi(\bar{x})$  is the (symmetric) Hessian of  $\varphi$  at  $\bar{x}$ . Sometimes the second-order construction (2.11) is called the “generalized Hessian” of  $\varphi$  at the reference point [39]. Note also that for the so-called  $\mathcal{C}^{1,1}$  functions (i.e., continuously differentiable ones with locally Lipschitzian derivatives around  $\bar{x}$ ; another notation is  $\mathcal{C}^{1+}$ ), we have the representation

$$\partial^2\varphi(\bar{x})(u) = \partial\langle u, \nabla\varphi\rangle(\bar{x}), \quad u \in \mathbb{R}^n,$$

via the basic first-order subdifferential (2.2) of the derivative scalarization  $\langle u, \nabla\varphi\rangle(x) := \langle u, \nabla\varphi(x) \rangle$  as  $x \in \mathbb{R}^n$ ; see [29, Proposition 1.120]. It is worth emphasizing that the second-order subdifferential (2.11) as well as the generating coderivative and first-order subdifferential mappings are *dual-space* intrinsically nonconvex constructions, which cannot correspond by duality to any derivative-like objects in primal spaces studied, e.g., in [2, 45].

Following the scheme of Definition 2.1 and keeping the coderivative (2.9) as the underlying element of our approach while using different first-order subdifferentials in (2.11), we may define a variety of second-order constructions of type (2.11). In particular, for functions  $\varphi: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$  of  $(x, w) \in \mathbb{R}^n \times \mathbb{R}^d$  there are two reasonable ways of introducing partial second-order subdifferentials; cf. [19]. To proceed, define the *partial first-order* subgradient mapping  $\partial_x\varphi: \mathbb{R}^n \times \mathbb{R}^d \rightrightarrows \mathbb{R}^n$  by

$$\partial_x\varphi(x, w) := \left\{ \text{set of subgradients } v \text{ of } \varphi_w := \varphi(\cdot, w) \text{ at } x \right\} = \partial\varphi_w(x).$$

Then given  $(\bar{x}, \bar{w})$  and  $\bar{y} \in \partial_x\varphi(\bar{x}, \bar{w})$ , define the *partial second-order subdifferential* of  $\varphi$  with respect to  $x$  of at  $(\bar{x}, \bar{w})$  relative to  $\bar{y}$  by

$$(2.12) \quad \partial_x^2\varphi(\bar{x}, \bar{w}, \bar{y})(u) := (D^*\partial\varphi_{\bar{w}})(\bar{x}, \bar{y})(u) = \partial^2\varphi_{\bar{w}}(\bar{x}, \bar{y})(u), \quad u \in \mathbb{R}^n,$$

with  $\varphi_{\bar{w}}(x) = f(x, \bar{w})$ . On the other hand, we can define the *extended partial second-order subdifferential* of  $\varphi$  with respect to  $x$  of at  $(\bar{x}, \bar{w})$  relative to  $\bar{y}$  by

$$(2.13) \quad \tilde{\partial}_x^2\varphi(\bar{x}, \bar{w}, \bar{y})(u) := (D^*\partial_x\varphi)(\bar{x}, \bar{w}, \bar{y})(u), \quad u \in \mathbb{R}^n.$$

As argued in [19], constructions (2.12) and (2.13) are not the same even in the case of  $\mathcal{C}^2$  functions when (2.12) reduces to  $\nabla_{xx}^2\varphi(\bar{x}, \bar{w})(u)$  while (2.13) comes out as  $(\nabla_{xx}^2\varphi(\bar{x}, \bar{w})u, \nabla_{xw}^2\varphi(\bar{x}, \bar{w})u)$ . This happens due to the involvement of  $w \rightarrow \bar{w}$  in the limiting procedure to define the extended partial second-order subdifferential set  $\tilde{\partial}_x^2\varphi(\bar{x}, \bar{w}, \bar{y})(u)$ , which is hence larger than (2.12). Note that both partial second-order constructions (2.12) and (2.13) are proved to be useful in applications; see, e.g., [18, 19] for more details.

It has been well recognized and documented (see, e.g., [3, 29, 30, 45, 46] and the references therein) that the first-order limiting constructions (2.2), (2.5), and (2.9) enjoy full calculi, which are crucial for their numerous applications. Based on definitions (2.11) of the second-order subdifferential and its partial counterparts (2.12) and (2.13), it is natural to try to combine calculus results for first-order subgradients with those for coderivatives to arrive at the corresponding second-order calculus rules. However, there are nontrivial complications to proceed in this way due to the fact that general results of the first-order subdifferential calculus hold as *inclusions* while the coderivative (2.9) does *not* possesses any *monotonicity* properties. Thus the initial requirement arises on selecting classes of functions for which calculus rules for first-order subgradients hold as *equalities*. Proceeding in this direction, a number of second-order calculus rules have been established in [18, 26, 28, 29, 31, 32, 34] for full while *not for partial* second-order subdifferentials in finite and infinite dimensions.

In the next section we obtain new second-order chain rules applied to full and partial second-order subdifferentials and develop, in particular, a direct approach based on quadratic penalties to derive general results for strongly amenable compositions.

### 3 Second-Order Subdifferential Chain Rules

Given a vector function  $h: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^m$  with  $m \leq n$  and a proper extended-real-valued function  $\theta: \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ , consider the composition

$$(3.1) \quad \varphi(x, w) = (\theta \circ h)(x, w) := \theta(h(x, w))$$

with  $x \in \mathbb{R}^n$  and  $w \in \mathbb{R}^d$ . Our first theorem provides *exact formulas* for calculating the partial second-order subdifferentials (2.12) and (2.13) of composition (3.1) under the *full rank condition* on the partial derivative (Jacobian matrix)  $\nabla_x h(\bar{x}, \bar{w})$  at the reference point.

**Theorem 3.1 (exact second-order chain rules with full rank condition).** *Given a point  $(\bar{x}, \bar{w}) \in \mathbb{R}^n \times \mathbb{R}^d$ , suppose that  $\theta$  in (3.1) is finite at  $\bar{z} := h(\bar{x}, \bar{w})$ , that  $h(\cdot, \bar{w}): \mathbb{R}^n \rightarrow \mathbb{R}^m$  is continuously differentiable around  $\bar{x}$  with the full row rank condition*

$$(3.2) \quad \text{rank } \nabla_x h(\bar{x}, \bar{w}) = m,$$

*and that the mapping  $\nabla_x h(\cdot, \bar{w}): \mathbb{R}^n \rightarrow \mathbb{R}^m$  is strictly differentiable at  $\bar{x}$ . Pick any  $\bar{y} \in \partial_x \varphi(\bar{x}, \bar{w})$  and denote by  $\bar{v}$  a unique vector satisfying the relationships*

$$\bar{v} \in \partial \theta(\bar{z}) \quad \text{and} \quad \nabla_x h(\bar{x}, \bar{w})^* \bar{v} = \bar{y}.$$

*Then we have the chain rule equality for the partial second-order subdifferential (2.12):*

$$(3.3) \quad \partial_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) = \nabla_{xx}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u + \nabla_x h(\bar{x}, \bar{w})^* \partial^2 \theta(\bar{z}, \bar{v})(\nabla_x h(\bar{x}, \bar{w})u), \quad u \in \mathbb{R}^n.$$

*If in addition the mapping  $h: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^m$  is continuously differentiable around  $(\bar{x}, \bar{w})$  and its derivative  $\nabla h: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^m$  is strictly differentiable at  $(\bar{x}, \bar{w})$ , then we have*

$$(3.4) \quad \begin{aligned} \tilde{\partial}_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) &= \left( \nabla_{xx}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u, \nabla_{xw}^2 \langle \bar{v}, h \rangle(\bar{x}, \bar{w})u \right) \\ &+ \left( \nabla_x h(\bar{x}, \bar{w})^* \partial^2 \theta(\bar{z}, \bar{v})(\nabla_x h(\bar{x}, \bar{w})u), \nabla_w h(\bar{x}, \bar{w})^* \partial^2 \theta(\bar{z}, \bar{v})(\nabla_x h(\bar{x}, \bar{w})u) \right) \end{aligned}$$

*whenever  $u \in \mathbb{R}^n$  for the extended partial second-order subdifferential (2.13).*

**Proof.** We derive the chain rule (3.4) for the extended partial second-order subdifferential; the proof of (3.3) is just a simplification of the one given below.

On the first-order subdifferential level we have from [29, Proposition 1.112] under the assumptions made (and from [45, Exercise 10.7] under some additional assumptions) that there is a neighborhood  $U$  of  $(\bar{x}, \bar{w})$  such that

$$\partial_x \varphi(x, w) = \left\{ y \in \mathbb{R}^n \mid \exists v \in \partial \theta(h(x, w)) \text{ with } \nabla_x h(x, w)^* v = y \right\} \quad \text{for all } (x, w) \in U.$$

For any fixed  $\bar{y} \in \partial_x \varphi(\bar{x}, \bar{w})$ , this gives us locally around  $(\bar{x}, \bar{w}, \bar{y})$  the graph representation

$$(3.5) \quad \text{gph } \partial_x \varphi = \left\{ (x, w, y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \mid \begin{array}{l} \exists (p, v) \in \text{gph } \theta \text{ such that} \\ h(x, w) = p, \nabla_x h(x, w)^* v = y \end{array} \right\}.$$

Consider now the two possible cases in the graph representation (3.5): **(i)** the “square” case when  $m = n$  and **(ii)** the “general” one when  $m < n$ .



In the square case (i) we have by the full rank condition (3.2) that the matrix  $\nabla_x h(x, w)$  is invertible for  $(x, w)$  near  $(\bar{x}, \bar{w})$ , and hence (3.5) can be rewritten as

$$(3.6) \quad \text{gph } \partial_x \varphi = \left\{ (x, w, y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \mid f(x, w, y) \in \text{gph } \partial \theta \right\}$$

via the mapping  $f: \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^{2n}$  given by

$$(3.7) \quad f(x, w, y) := \left( h(x, w), (\nabla_x h(x, w))^* y \right) \text{ for } (x, w, y) \text{ near } (\bar{x}, \bar{w}, \bar{y}).$$

In other words, representation (3.6) can be expressed via the preimage/inverse image of the set  $\text{gph } \partial \theta$  under the mapping  $f$  as follows:

$$(3.8) \quad \text{gph } \partial_x \varphi = f^{-1}(\text{gph } \partial \theta).$$

Since  $\nabla_x h$  is assumed to be strictly differentiable at  $(\bar{x}, \bar{w})$ , the mapping  $f$  in (3.7) is strictly differentiable at  $(\bar{x}, \bar{w}, \bar{y})$  and, by (3.2) with  $m = n$ , its Jacobian matrix  $\nabla f(\bar{x}, \bar{w}, \bar{y})$  has full row rank  $2n$ . Employing [29, Theorem 1.17] to (3.8) gives us

$$(3.9) \quad N((\bar{x}, \bar{w}, \bar{y}); \text{gph } \partial_x \varphi) = \nabla f(\bar{x}, \bar{w}, \bar{y})^* N(f(\bar{x}, \bar{w}, \bar{y}); \text{gph } \partial \theta).$$

Now we calculate the derivative/Jacobian matrix of  $f$  at  $(\bar{x}, \bar{w}, \bar{y})$  by using the particular structure of  $f$  in (3.7), the classical chain rule, and the well-known Leach inverse function theorem for strictly differentiable mappings; see, e.g., [10, 29]. Define the mappings  $f_1: \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $f_2: \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  with  $f = (f_1, f_2)$  by  $f_1(x, w, y) := f(x, w)$  and

$$f_2(x, w, y) := (\nabla_x h(x, w))^{-1} y \text{ for } (x, w, y) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n.$$

It is clear that  $\nabla f_1(\bar{x}, \bar{w}, \bar{y}) = (\nabla f(\bar{x}, \bar{w}), 0)$ , while for calculating  $\nabla f_2(\bar{x}, \bar{w}, \bar{y})$  we introduce two auxiliary mapping  $g: \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $q: \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}$  by

$$g(x, w, p) := \nabla_x h(x, w)^* p \text{ and } q(x, w, p) := \langle p, h(x, w) \rangle \text{ for } (x, w, p) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n.$$

Note that  $g(x, w, p) = \nabla q_x(x, w, p)^*$  and that  $g(x, w, f_2(x, w, y)) - v = 0$ . Differentiating the latter equality gives us

$$(3.10) \quad \nabla_x g(x, w, f_2(x, w, y)) + \nabla_p g(x, w, f_2(x, w, y)) \nabla_x f_2(x, w, y) = 0.$$

Observing further that  $\nabla_x g(x, w, p) = \nabla_x (\nabla_x q(x, w, p))^* = \nabla_{xx}^2 q(x, w, p)$ , we get from (3.10) and the definitions above that the equation

$$\nabla_{xx}^2 \langle (\nabla_x h(x, w))^{-1} y, h(x, w) \rangle + \nabla_x h(x, w)^* \nabla f_2(x, w, y) = 0$$

is satisfied, which implies in turn the representation of the partial derivative

$$(3.11) \quad \nabla_x f_2(x, w, y) = -(\nabla_x h(x, w))^{-1} \cdot \nabla_{xx}^2 \langle (\nabla_x h(x, w))^{-1} y, h(x, w) \rangle.$$

Similarly we have the following representation of the other partial derivative of  $f_2$ :

$$(3.12) \quad \nabla_w f_2(x, w, y) = -(\nabla_x h(x, w))^{-1} \cdot \nabla_{xw}^2 \langle (\nabla_x h(x, w))^{-1} y, h(x, w) \rangle.$$

Taking into account that  $\nabla_y f_2(x, w, y) = (\nabla_x h(x, w))^{-1}$  gives us finally

$$(3.13) \quad \nabla f(\bar{x}, \bar{w}, \bar{y}) = \begin{bmatrix} \nabla_x h(\bar{x}, \bar{w}) & \nabla_w h(\bar{x}, \bar{w}) & 0 \\ \nabla_x f_2(\bar{x}, \bar{w}, \bar{y}) & \nabla_w f_2(\bar{x}, \bar{w}, \bar{y}) & (\nabla_x h(\bar{x}, \bar{w}))^{-1} \end{bmatrix},$$

where  $\nabla_x f_2(\bar{x}, \bar{w}, \bar{y})$  and  $\nabla_w f_2(\bar{x}, \bar{w}, \bar{y})$  are calculated in (3.11) and (3.12), respectively. The second-order chain rule (3.4) in the square case (i) follows now from substituting (3.13) into (3.9) and by using then the definitions of the constructions involved and elementary transformations.

It remains to consider the general case (ii) with  $m < n$ . This case can be reduced to the previous one by introducing a linear mapping  $\bar{h}: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^{n-m}$  such that the mapping

$$\bar{h}(x, w) := (h(x, w), \tilde{h}(x, w)) \text{ from } \mathbb{R}^n \times \mathbb{R}^d \text{ to } \mathbb{R}^n$$

has full rank. It can be done, e.g., by choosing a basis  $\{a_1, \dots, a_{n-m}\}$  for the  $n - m$ -dimensional spaces  $\{u \in \mathbb{R}^n \mid \nabla h_x(\bar{x}, \bar{w})u = 0\}$  and letting  $\tilde{h}(x, w) := (h(x, w), \langle a_1, x \rangle, \dots, \langle a_{n-m}, x \rangle)$ ; cf. [45, Exercise 6.7] for a first-order setting. Then viewing  $\varphi$  as  $\bar{\theta} \circ \bar{h}$  with  $\bar{\theta}(z, p) := \theta(z)$  for all  $z \in \mathbb{R}^m$  and  $p \in \mathbb{R}^{n-m}$  reduces (ii) to (i) and thus completes the proof of the theorem.  $\triangle$

Some remarks on the results related to those obtained in Theorem 3.1 are in order.

**Remark 3.2 (discussions on second-order chain rules with full rank/surjectivity conditions).** Previously known second-order chain rules of type (3.3) were derived for the full second-order subdifferential (2.11), where condition (3.2) was written as  $\text{rank} \nabla h(\bar{x}) = m$ . To the best of our knowledge, the first result in this direction was obtained in [31, Theorem 3.4] with the inclusion “ $\subset$ ” in (3.3). Various infinite-dimensional extensions of (3.3) in the inclusion and equality forms were derived in [28, 34] and [29, Theorem 1.127] with imposing the *surjectivity condition* on the derivative  $\nabla h(\bar{x})$  as the counterpart of (3.2) in infinite-dimensional spaces. Observe that the proof of (3.4) given above in case (i) corresponding to the invertible partial derivative  $\nabla_x h(\bar{x}, \bar{w})$  holds in any Banach space, while the device in case (ii) is finite-dimensional.

Next we explore the possibility of deriving second-order chain rules for (3.1) when the rank condition (3.2) may not be satisfied. This can be done for broad classes of amenable functions defined in the way originated in [37], which are overwhelmingly encountered in finite-dimensional parametric optimization. Recall [20] that a proper function  $\varphi: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$  is *strongly amenable* in  $x$  at  $\bar{x}$  with *compatible parametrization* in  $w$  at  $\bar{w}$  if there is a neighborhood  $V$  of  $(\bar{x}, \bar{w})$  on which  $\varphi$  is represented in the composition form (3.1), where  $h$  is of class  $\mathcal{C}^2$  while  $\theta$  is a proper, l.s.c., convex function such that the first-order qualification condition

$$(3.14) \quad \partial^\infty \theta(h(\bar{x}, \bar{w})) \cap \ker \nabla_x h(\bar{x}, \bar{w})^* = \{0\}$$

involving the singular subdifferential (2.3) is satisfied. The latter qualification condition automatically holds if either  $\theta$  is locally Lipschitzian around  $h(\bar{x}, \bar{w})$  or the full rank condition (3.2) is fulfilled, since it is equivalent to

$$\ker \nabla_x h(\bar{x}, \bar{w})^* := \{v \in \mathbb{R}^n \mid 0 = \nabla_x h(\bar{x}, \bar{w})^* v\} = \{0\}.$$

Properties of strongly amenable compositions  $\varphi(x) = \theta(h(x))$  and related functions are largely investigated in [37, 38, 45]; most of them hold also for strongly amenable compositions (3.1) with compatible parametrization [19, 20]. Strong amenability is a property that bridges between smoothness and convexity while covering at the same time a great many of the functions that are of interest as the essential objective in minimization problems; see [45] for more details.

The next theorem establishes second-order subdifferential chain rules of the inclusion type for strongly amenable compositions with no full rank requirement (3.2).

**Theorem 3.3 (second-order chain rules for strongly amenable compositions).** *Let  $\varphi: \mathbb{R}^n \times \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$  be strongly amenable in  $x$  at  $\bar{x}$  with compatible parametrization in  $w$  at  $\bar{w}$ , and let  $\bar{y} \in \partial_x \varphi(\bar{x}, \bar{w})$ . Denote  $\bar{z} := h(\bar{x}, \bar{w})$  and consider the nonempty set*

$$M(\bar{x}, \bar{w}, \bar{y}) := \left\{ v \in \mathbb{R}^m \mid v \in \partial \theta(\bar{z}) \text{ with } \nabla_x h(\bar{x}, \bar{w})^* v = \bar{y} \right\}.$$

Assume the fulfillment of the second-order qualification condition:

$$(3.15) \quad \partial^2\theta(\bar{z}; v)(0) \cap \ker \nabla_x h(\bar{x}, \bar{w})^* = \{0\} \text{ for all } v \in M(\bar{x}, \bar{w}, \bar{y}).$$

Then we have the following chain rules for the partial second-order subdifferentials (2.12) and (2.13), respectively, valued for all  $u \in \mathbb{R}^n$ :

$$(3.16) \quad \partial_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) \subset \bigcup_{v \in M(\bar{x}, \bar{w}, \bar{y})} \nabla_{xx}^2 \langle v, h \rangle(\bar{x}, \bar{w})u + \nabla_x h(\bar{x}, \bar{w})^* \partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u),$$

$$(3.17) \quad \begin{aligned} \tilde{\partial}_x^2 \varphi(\bar{x}, \bar{w}, \bar{y})(u) &\subset \bigcup_{v \in M(\bar{x}, \bar{w}, \bar{y})} \left( \nabla_{xx}^2 \langle v, h \rangle(\bar{x}, \bar{w})u, \nabla_{xw}^2 \langle v, h \rangle(\bar{x}, \bar{w})u \right) \\ &+ \left( \nabla_x h(\bar{x}, \bar{w})^* \partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u), \nabla_w h(\bar{x}, \bar{w})^* \partial^2\theta(\bar{z}, v)(\nabla_x h(\bar{x}, \bar{w})u) \right). \end{aligned}$$

**Proof.** For brevity and simplicity of the arguments and notation, we present a detailed proof just for the full second-order subdifferential (2.11) of the strongly amenable nonparameterized compositions  $\varphi(x) = \theta(h(x))$  in which case both formulas (3.16) and (3.17) reduce to

$$(3.18) \quad \partial^2 \varphi(\bar{x}, \bar{y})(u) \subset \bigcup_{\substack{v \in \partial\theta(\bar{z}) \\ \nabla h(\bar{x})^* v = \bar{y}}} \left( \nabla^2 \langle v, h \rangle(\bar{x})u + \nabla h(\bar{x})^* \partial^2\theta(\bar{z}, v)(\nabla h(\bar{x})u) \right)$$

with  $\bar{z} = h(\bar{x})$  under the *basic second-order qualification condition*

$$(3.19) \quad \partial^2\theta(\bar{z}; v)(0) \cap \ker \nabla h(\bar{x})^* = \{0\} \text{ whenever } v \in \partial\theta(\bar{z}) \text{ and } \nabla h(\bar{x})^* v = \bar{y}.$$

The reader can readily check that the method of quadratic penalties developed below perfectly works for the case of partial second-order subdifferentials to produce the chain rule inclusions (3.16) and (3.17) under the “partial” second-order qualification condition (3.15).

We begin with observing that the first-order chain rule

$$(3.20) \quad \partial\varphi(x) = \nabla h(x)^* \partial\theta(h(x)) \text{ whenever } x \in U$$

holds as equality for strongly amenable compositions on some neighborhood  $U$  of  $\bar{x}$ . Indeed, it follows from the more general chain rule of [29, Theorem 3.41(iii)] due to (2.10) and the particular properties of strongly amenable functions summarized in [45, Exercise 10.25].

Now we proceed with calculating of the second-order subdifferential  $\partial^2 \varphi(\bar{x}, \bar{y})$  for the given first-order subgradient  $\bar{y} \in \partial\varphi(\bar{x})$ . The definitions in (2.1), (2.9), and (2.5) suggest to us calculating the regular normal cone  $\hat{N}((x, y); \Omega)$  to the subdifferential graph  $\Omega := \text{gph } \partial\varphi$  of  $\varphi$  at points  $(x, y) \in \text{gph } \partial\varphi$  near  $(\bar{x}, \bar{y})$  and then passing to the limit therein as  $(x, y) \rightarrow (\bar{x}, \bar{y})$ . To simplify notation, let us focus first on calculating  $\hat{N}((\bar{x}, \bar{y}); \Omega)$  for the graphical set  $\Omega$ . Developing a *variational approach* to subdifferential calculus and employing the *smooth variational description* of regular normals from [45, Theorem 6.11] and [29, Theorem 1.30], we have that  $(\omega, -\xi) \in \hat{N}((\bar{x}, \bar{y}); \Omega)$  if and only if there is a smooth function  $\vartheta: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$(3.21) \quad \text{argmin}_{(x, y) \in \Omega} \vartheta(x, y) = \{(\bar{x}, \bar{y})\} \text{ and } \nabla \vartheta(\bar{x}, \bar{y}) = (-\omega, \xi).$$

Using the first-order chain rule formula (3.20) allows us to transform the minimization problem in (3.21) into the following one:

$$(3.22) \quad \begin{cases} \text{minimize } \vartheta(x, \nabla h(x)^* v) & \text{over all} \\ x \in \mathbb{R}^n, (v, z) \in \text{gph } \partial\theta & \text{with } h(x) - z = 0. \end{cases}$$

We know from (3.21) that  $(x, z, v)$  is an optimal solution to (3.22) if and only if

$$(3.23) \quad x = \bar{x}, \quad z = \bar{z}, \quad \text{and} \quad \nabla h(\bar{x})^* v = \bar{y}.$$

Let  $G := \text{gph } \theta$  and for any  $\varepsilon > 0$  consider the *quadratic penalty problem*:

$$(3.24) \quad \begin{cases} \text{minimize } \vartheta(x, \nabla h(x)^* v) + \frac{1}{2\varepsilon} \|h(x) - z\|^2 \\ \text{over all } x \in \mathbb{R}^n \text{ and } (z, v) \in G. \end{cases}$$

Denoting by  $\varepsilon \mapsto S(\varepsilon)$  the optimal solution map for problem (3.24), observe that it is of *closed graph* and uniformly *bounded* around  $\varepsilon = 0$ . Indeed, it follows from the closedness of  $\text{gph } \partial\theta$  due to the convexity and lower semicontinuity of  $\theta$ , the local continuity of  $\vartheta$  and  $\nabla h$ , and the uniform boundedness near  $(\bar{x}, \bar{y})$  of the mapping  $M: \Omega \rightrightarrows \mathbb{R}^m$  given by

$$(3.25) \quad M(x, y) := \left\{ v \in \partial\theta(h(x)) \mid \nabla h(x)^* v = y \right\}$$

with  $\Omega = \text{gph } \partial\varphi$ . The latter local boundedness can be easily verified arguing by contradiction due the qualification condition (3.14) in the definition of amenable functions.

Now consider a sequence of values  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow \infty$  and pick a triple  $(x_k, z_k, v_k) \in S(\varepsilon_k)$  for all  $k \in \mathbb{N}$ . By the local boundedness of the solution map  $S(\varepsilon)$  near  $\varepsilon = 0$  the sequence  $\{v_k\}$  is bounded, and thus it has a cluster point  $\bar{v}$ . Without loss of generality we suppose that  $v_k \rightarrow \bar{v}$  as  $k \rightarrow \infty$  and get therefore that

$$(3.26) \quad x_k \rightarrow \bar{x}, \quad z_k := h(x_k) \rightarrow \bar{z}, \quad \text{and} \quad y_k := \nabla h(x_k)^* v_k \rightarrow \nabla h(\bar{x})^* \bar{v} = \bar{y}$$

with  $(\bar{z}, \bar{v}) \in G$ . By the observation above (3.23), the triple  $(\bar{x}, \bar{z}, \bar{v})$  is the unique optimal solution to the unperturbed problem (3.22). On the other hand, applying the first-order necessary optimality conditions from [45, Theorem 6.12] to the solution  $(x_k, z_k, v_k)$  of the penalized problem (3.24) with a smooth cost function and a geometric constraint gives us

$$\begin{aligned} \nabla_x \left[ \vartheta(x, \nabla h(z)^* v_k) + \frac{1}{\varepsilon_k} \|h(x) - z_k\|^2 \right] \Big|_{x=x_k} &= 0, \\ -\nabla_{z,v} \left[ \vartheta(x_k, \nabla h(x_k)^* v) + \frac{1}{2\varepsilon_k} \|h(x_k) - z\|^2 \right] \Big|_{(z,v)=(z_k,v_k)} &\in \widehat{N}((z_k, v_k); G) \end{aligned}$$

for all  $k \in \mathbb{N}$ . Denoting  $p_k := [h(x_k) - z_k]/\varepsilon_k$ , these conditions calculate out to

$$(3.27) \quad \nabla_x \vartheta(x_k, y_k) + \nabla^2 \langle v_k, h \rangle(x_k) \nabla_y \vartheta(x_k, y_k) + \nabla \langle p_k, h \rangle(x_k) = 0,$$

$$(3.28) \quad \left( p_k, -\nabla h(x_k) \nabla_y \vartheta(x_k, y_k) \right) \in \widehat{N}((z_k, v_k); G).$$

By passing above to subsequences as  $k \rightarrow \infty$  if needed, we can reduce the situation to considering one of the following two cases:

**Case 1:**  $\{p_k\}$  converges to some  $\bar{p}$ .

**Case 2:**  $p_k \rightarrow \infty$  while  $\{p_k/\|p_k\|\}$  converges to some  $\bar{p} \neq 0$ .

In Case 1 it follows from (3.26), (3.27), and (3.28) that

$$(3.29) \quad \nabla_x \vartheta(\bar{x}, \bar{y}) + \nabla^2 \langle \bar{v}, h \rangle(\bar{x}) \nabla_y \vartheta(\bar{x}, \bar{y}) + \nabla \langle \bar{p}, h \rangle(\bar{x}) = 0,$$

$$(3.30) \quad \left( \bar{p}, -\nabla h(\bar{x}) \nabla_y \vartheta(\bar{x}, \bar{y}) \right) \in \widehat{N}((\bar{z}, \bar{v}); G),$$

where  $\nabla_x \vartheta(\bar{x}, \bar{y}) = -\omega$  and  $\nabla_y \vartheta(\bar{x}, \bar{y}) = \xi$  by the second equality in (3.21).

In Case 2 we get, dividing first both parts of (3.27) and (3.28) by  $\|p_k\|$  and then passing to the limit therein as  $k \rightarrow \infty$ , that

$$(3.31) \quad \nabla h(\bar{x})^* \bar{p} = 0 \quad \text{and} \quad (\bar{p}, 0) \in \widehat{N}((\bar{z}, \bar{v}); G) \quad \text{with} \quad \|\bar{p}\| = 1.$$

Thus, by talking into account our choice of  $(\omega, -\xi) \in \widehat{N}((\bar{x}, \bar{y}); \Omega)$  and the construction of  $M$  in (3.25), we deduce from (3.29)–(3.31) the existence of  $\bar{v} \in M(\bar{x}, \bar{y})$  and  $\bar{p}$  satisfying either (3.29) and (3.30) or (3.31). Since the arguments above equally hold for every point  $(x, y) \in \text{gph } \partial\varphi$  near  $(\bar{x}, \bar{y})$ , they ensure the following description of the regular normal cone to  $\Omega = \text{gph } \varphi$  at points  $(x, y) \in \Omega$  in a neighborhood of the reference one  $(\bar{x}, \bar{y})$ , where  $G = \text{gph } \partial\theta$ :

$$(3.32) \quad \left[ \begin{array}{l} (\omega, -\xi) \in \widehat{N}((x, y); \Omega) \implies \exists v \in M(x, y), p \in \mathbb{R}^m \text{ such that} \\ \text{either: } \begin{cases} \omega = \nabla^2 \langle v, h \rangle(x) \xi + \nabla h(x)^* p \\ \text{with } (p, -\nabla h(x) \xi) \in \widehat{N}(h(x), v); G \end{cases} \\ \text{or: } \nabla h(x)^* p = 0 \text{ with } (p, 0) \in \widehat{N}(h(x), v); G, \quad \|p\| = 1. \end{array} \right.$$

Next we take any basic normal  $(\omega, -\xi) \in N(\bar{x}, \bar{y}); \Omega$  (not just a regular one) and by (2.5) find sequences  $(x_k, y_k) \rightarrow (\bar{x}, \bar{y})$  and  $(\omega_k, -\xi_k) \rightarrow (\omega, -\xi)$  as  $k \rightarrow \infty$  satisfying

$$(x_k, y_k) \in \Omega \quad \text{and} \quad (\omega_k, -\xi_k) \in \widehat{N}((x_k, y_k); \Omega) \quad \text{for all } k \in \mathbb{N}.$$

Employing the description of regular normals (3.32) ensures the existence of  $v_k \in M(x_k, y_k)$  and  $p_k \in \mathbb{R}^m$  such that the either/or alternative in (3.32) holds for each  $k \in \mathbb{N}$ . Due to the established local boundedness of the mapping  $M$ , suppose with no loss of generality that

$$v_k \rightarrow v \quad \text{as } k \rightarrow \infty \quad \text{for some } v \in M(\bar{x}, \bar{y}).$$

By a further passage to subsequences, we can reduce the situation to where just one of the "either/or" parts of the alternative in (3.32) holds for all  $k$ . Consider first the "or" part of this alternative, i.e., the validity of

$$\nabla h(x_k)^* p_k = 0 \quad \text{with} \quad (p_k, 0) \in \widehat{N}(h(x_k), v_k); G, \quad \|p_k\| = 1 \quad \text{for all } k.$$

In this case the sequence  $\{p_k\}$  has a cluster point  $p$ , and thus we get

$$(3.33) \quad \nabla h(\bar{x})^* p = 0 \quad \text{with} \quad (p, 0) \in N((h(\bar{x}), v); G), \quad \|p\| = 1, \quad \text{and} \quad v \in M(\bar{x}, \bar{y})$$

by passing to the limit as  $k \rightarrow \infty$  and taking into account (2.5) and the continuity of  $f$  and  $\nabla h$ .

When the "either" part holds, we proceed similarly to Cases 1 and 2 above. In the first case there is  $p \in \mathbb{R}^m$  such that  $p_k \rightarrow p$ . Then the passage to the limit in

$$(3.34) \quad \omega_k = \nabla^2 \langle v_k, h \rangle(x_k) \xi_k + \nabla h(x_k)^* p_k, \quad (p_k, -\nabla h(x_k) \xi_k) \in \widehat{N}((h(x_k), v_k); G)$$

with taking into account the continuity assumptions and the convergence above, leads to

$$(3.35) \quad \omega = \nabla^2 \langle v, h \rangle(\bar{x}) \xi + \nabla h(\bar{x})^* p, \quad (p, -\nabla h(\bar{x}) \xi) \in \widehat{N}((h(\bar{x}), v); G).$$

In the remaining case we have  $\|p_k\| \rightarrow \infty$  and thus find  $p$  such that  $p_k/\|p_k\| \rightarrow p$  with  $\|p\| = 1$ . Divide now both sides of (3.34) by  $\|p_k\|$  for any large  $k$  and take the limit therein as  $k \rightarrow \infty$ .

Then we again arrive at (3.33). Unifying (3.33) and (3.35) gives as the description of basic normals to  $\Omega = \text{gph } \partial\varphi$  via the following alternative:

$$(3.36) \quad \left[ \begin{array}{l} (\omega, -\xi) \in N((\bar{x}, \bar{y}); \Omega) \implies \exists v \in M(\bar{x}, \bar{y}), p \in \mathbb{R}^m \text{ such that} \\ \text{either: } \left\{ \begin{array}{l} \omega = \nabla^2 \langle v, h \rangle(\bar{x})\xi + \nabla h(\bar{x})^* p \\ \text{with } (p, -\nabla h(\bar{x})\xi) \in N((h(\bar{x}), v); G) \end{array} \right. \\ \text{or: } \nabla h(\bar{x})^* p = 0 \text{ with } (p, 0) \in N(h(\bar{x}), v); G, \|p\| = 1. \end{array} \right.$$

Remembering the notation introduced in the theorem and the definitions of the constructions used, we see that the “either” part of (3.36) amounts to the second-order subdifferential inclusion (3.18) while the “or” part of (3.36) means the negation of the basic second-order qualification condition (3.19). Thus the assumed fulfillment of (3.19) shows that the “or” part of (3.36) does not hold, which justifies the validity of the second-order chain rule (3.18).

Repeating finally the arguments above with taking in to account that the partial counterparts of the first-order chain rule equality (3.20) are satisfied due to the results of [19, Proposition 3.4] (see also [45, Corollary 10.11] and [29, Corollary 3.44] in more general settings), we get the partial second-order subdifferential chain rules (3.16) and (3.17) under the partial second-order qualification condition (3.15).  $\triangle$

**Remark 3.4 (second-order chain rules with inclusions).** A chain rule in form (3.16) for the *full* second-order subgradient sets of strongly amenable compositions with compatible parametrization in finite dimensions was derived in [18] under the second-order subdifferential condition of type (3.15) with  $\nabla h(\bar{x}, \bar{w})$  replacing  $\nabla_x h(\bar{x}, \bar{w})$ . The proof in [18] was based on applying a coderivative chain rule to full first-order subdifferential mappings. A similar approach was employed in [28] and [29, Theorem 3.74] to derive second-order chain rules of type (3.18) in infinite dimensions under an appropriate infinite-dimensional counterpart of the second-order qualification condition (3.19). Although the results of [28, 29] are applied to a more general class of subdifferential regular functions  $\theta$  in (3.18), they require a number of additional assumptions in both finite and infinite dimensions. Finally, we mention a second-order chain rule of the inclusion type (3.18) obtained in [32, Theorem 3.1] for a special kind of strongly amenable compositions with the indicator function  $\theta = \delta(\cdot; \Theta)$  of a set  $\Theta$  in finite dimensions, which does not generally require the fulfillment of the second-order qualification condition (3.19) while imposing instead of a certain *calmness* assumption on some auxiliary multifunction. The latter holds, in particular, in the case of polyhedral sets  $\Theta$  due to seminal results of [42].

Next we show that the second-order chain rule formula (3.18), and hence those in (3.16) and (3.17), cannot be generally used for *precise calculating* the second-order subdifferentials of strongly amenable compositions: the inclusion therein may be *strict* even for fairly simple functions  $\theta$  and  $h$  in  $\varphi = \theta \circ h$  without a kind of full rank condition.

**Example 3.5 (strict inclusion in the second-order chain rule formula).** *The inclusion in (3.18) can be strict even when  $h$  is linear while  $\theta$  is piecewise linear and convex. Moreover, the set on right-hand side of (3.18) can be nonempty when the one on the left-hand side is empty.*

**Proof.** Let the functions  $h: \mathbb{R}^2 \rightarrow \mathbb{R}^4$  and  $\theta: \mathbb{R}^4 \rightarrow \mathbb{R}$  be given by

$$h(x_1, x_2) := (x_1, -x_1, x_2, -x_2) = Ax \text{ with } A := \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix},$$

$$\theta(z_1, z_2, z_3, z_4) := \max \left\{ z_1, z_2, z_3, z_4 \right\} = \sigma_M(z),$$

where  $M := \{v = (v_1, v_2, v_3, v_4) \in \mathbb{R}^4 \mid v_i \geq 0, \sum_{i=1}^4 v_i = 1\}$  is the unit simplex in  $\mathbb{R}^4$ , and where  $\sigma_\Omega$  stands for the support function of the set  $\Omega$ . Considering the composition  $\varphi(x) := \theta(h(x))$  on  $\mathbb{R}^2$ , observe that it can be represented as

$$(3.37) \quad \varphi(x) = \sigma_B(x) \text{ for } B := \left\{ (y_1, y_2) \in \mathbb{R}^2 \mid |y_1| + |y_2| \leq 1 \right\}.$$

Note that the outer function  $\theta$  in the strongly amenable (in fact fully amenable) composition  $\theta \circ h$  is *convex piecewise linear* in terminology of [45]; it can be equivalently described by [45, Theorem 2.49] as a function with the polyhedral epigraph.

Using the explicit form (3.37) of  $\varphi$  allows us to compute its second-order subdifferential  $\partial^2 \varphi(\bar{x}, \bar{y})$  with  $\bar{x} = (0, 0)$  and  $\bar{y} = (0, 0)$  directly by Definition 2.1. Indeed, we get from (3.37) that  $\partial \varphi(\bar{x}) = B$ , and hence  $\bar{y} \in \text{int } \partial \varphi(\bar{x})$ . This tells us that

$$(3.38) \quad \partial^2 \varphi(\bar{x}, \bar{y})(u) = \begin{cases} \mathbb{R}^2 & \text{if } u = (0, 0), \\ \emptyset & \text{if } u \neq (0, 0). \end{cases}$$

On the other hand, formula (3.18) reads as the inclusion  $\partial^2 \varphi(\bar{x}, \bar{y})(u) \subset Q(u)$  with

$$Q(u) := \bigcup \left\{ A^* \partial^2 \theta(0, v)(Au) \mid v \in M, A^* v = 0 \right\}.$$

Take  $\bar{u} = (0, 1)$ ,  $\bar{v} = (1/2, 1/2, 0, 0)$  and then check that  $A\bar{u} = (0, 0, 1, -1)$ ,  $\bar{v} \in M$ , and  $A^* \bar{v} = 0$ . This ensures the converse inclusion

$$Q(\bar{u}) \supset A^* \partial^2 \theta(0, \bar{v})(A\bar{u}),$$

which shows that we have  $Q(\bar{u}) \neq \emptyset$  provided that  $\partial^2 \theta(0, \bar{v})(A\bar{u}) \neq \emptyset$ . To check the latter, recall the representation of the outer function  $\theta = \sigma_M = \delta_M^*$  via the indicator function  $\delta_M = \delta(\cdot; M)$  of  $M$ . Thus we get the description

$$\omega \in \partial^2 \theta(0, \bar{v})(A\bar{u}) \iff -A\bar{u} \in \partial^2 \delta_M(\bar{v}, 0)(-\omega).$$

Since the set  $M$  is a convex polyhedron, an exact formula for  $\partial^2 \delta_M(0, \bar{v})$  is available from [9].

In order to state this formula, we need to deal with the *critical cone* for a convex polyhedron  $\Omega$  at  $x \in \Omega$  with respect to  $p \in \partial \delta_\Omega(x) = N_\Omega(x)$ ; this is a polyhedral cone defined by

$$K(x, p) := \{w \in T_\Omega(\bar{x}) \mid w \perp p\},$$

where  $T_\Omega(x)$  is the tangent cone (2.7) to  $\Omega$  at  $x$ . Recall that a *closed face*  $C$  of a polyhedral cone  $K$  is a polyhedral cone of the form

$$C := \{x \in K \mid x \perp v\} \text{ for some } v \in K^*,$$

where  $K^*$  denotes the polar of the cone  $K$ . By the proof of [9, Theorem 2] (see also [38, Proposition 4.4]) we have the following description of the second-order subdifferential of the indicator function for a convex polyhedron:

$$(3.39) \quad w \in \partial^2 \delta_\Omega(x, p)(u) \iff \begin{cases} \text{there exist closed faces } C_1 \subset C_2 \text{ of } K(x, p) \\ \text{with } u \in C_1 - C_2, w \in (C_2 - C_1)^*. \end{cases}$$

Applying this to our setting with the simplex  $\Omega = M$ , we get the critical cone

$$K = T_M(\bar{v}) \cap 0^\perp = T_M(\bar{v}) = \left\{ (\omega_1, \omega_2, \omega_3, \omega_4) \mid \begin{array}{l} \omega_1 + \omega_2 = 0 \\ \omega_3 \geq 0, \omega_4 \geq 0 \end{array} \right\}.$$

It follows from the second-order subdifferential formula (3.39) that

$$-A\bar{u} \in \partial^2 \delta_M(\bar{v}, 0)(-\omega) \iff \begin{cases} \text{there exist closed faces } C_1 \subset C_2 \text{ of } K \\ \text{with } \omega \in C_2 - C_1, -A\bar{u} \in (C_2 - C_1)^*. \end{cases}$$

Observe that the closed faces of  $K$  have the form

$$\{(\omega_1, \omega_2) \mid \omega_1 + \omega_2 = 0\} \times I \times J, \text{ where } I, J \text{ can be either } \mathbb{R}_+ \text{ or } \{0\}.$$

Denoting by  $L$  the subspace in  $\{\dots\}$  of the formula above, we have the following possibilities:

$$C_1 - C_2 = L \times (I_1 - I_2) \times (J_1 - J_2) \text{ and } (C_1 - C_2)^* = L^\perp \times (I_1 - I_2)^* \times (J_1 - J_2)^*,$$

where  $I_1 - I_2$  and  $J_1 - J_2$  can be  $\mathbb{R}$ ,  $\mathbb{R}_+$ , and  $\{0\}$  while, respectively,  $(I_1 - I_2)^*$  and  $(J_1 - J_2)^*$  can be  $\{0\}$ ,  $\mathbb{R}_-$ , and  $\mathbb{R}$ . Setting now  $(I_1 - I_2)^* = \mathbb{R}_-$  and  $(J_1 - J_2)^* = \mathbb{R}$ , we get

$$-A\bar{u} \in \partial^2 \delta_M(\bar{v}, 0)(-\omega), \text{ or equivalently } \omega \in \partial^2 \theta(0, \bar{v})(A\bar{u})$$

whenever  $\omega \in L \times \mathbb{R}_+ \times \{0\}$ . Taking, e.g.,  $\omega = (0, 0, 1, 0)$  gives us  $\partial^2 \theta(0, \bar{v})(A\bar{u}) \neq \emptyset$ . Hence the set  $Q(\bar{u})$  on the right-hand side of (3.18) is nonempty while  $\partial^2 \varphi(\bar{x}, \bar{y})(\bar{u}) = \emptyset$  by (3.38).  $\triangle$

It is not hard to check that the second-order qualification condition (3.19) does not hold in Example 3.5. Thus, besides the emphasis above, this example can be considered as a counterexample to equality in the second-order chain rule (3.18) with no full rank condition on the derivative and also as an illustration of the possible validity of the inclusion in (3.18) without the second-order qualification condition (3.19). In the next section we show that the fulfillment of (3.19) yields, for a large class of outer functions  $\theta$  in amenable compositions  $\theta \circ h$  including the one in Example 3.5, that the full rank condition must be satisfied, and so the second-order chain rule formula (3.18) holds in fact as equality.

## 4 Analysis of the Basic Second-Order Qualification Condition and Calculating Second-Order Subdifferentials

This section is mainly devoted to analyzing the basic second-order qualification condition (3.19); our analysis equally works for the partial second-order counterpart (3.15).

It is clear that (3.19) holds under the full rank condition

$$(4.1) \quad \text{rank } \nabla h(\bar{x}) = m \iff \ker \nabla h(\bar{x})^* = \{0\},$$

which ensures the equality in (3.18) with a unique vector  $v$  in (3.18) and (3.19) by Theorem 3.1. Another setting where condition (3.19) automatically holds is when  $\theta$  is of class  $\mathcal{C}^{1,1}$  around  $\bar{z} = h(\bar{x})$ , which however excludes the case of extended-real-valued outer functions typical in applications of amenable compositions in variational analysis and optimization.

We show in what follows that, for large classes of amenable compositions, the second-order qualification condition (3.19) yields in fact the validity of the full rank condition (4.1) and hence the exact second-order subdifferential chain rule formula

$$(4.2) \quad \partial^2 \varphi(\bar{x}, \bar{y})(u) = \left( \nabla^2 \langle \bar{v}, h \rangle(\bar{x})u + \nabla h(\bar{x})^* \partial^2 \theta(\bar{z}, \bar{v})(\nabla h(\bar{x})u) \right) \text{ for all } u \in \mathbb{R}^n$$

with a unique vector  $\bar{v} \in \partial \theta(\bar{z})$  satisfying  $\nabla h(\bar{x})^* v = \bar{y}$ . As a by-product of our approach, we efficiently calculate the second-order subdifferential of functions belonging to such classes.



Recall [37, 45] that a strongly amenable function  $\varphi$  is *fully amenable* at  $\bar{x}$  if the outer function  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  in its composite representation  $\varphi = \theta \circ h$  can be chosen as *piecewise linear quadratic*. The latter class includes piecewise linear functions discussed in Example 3.5 and have their domains as well as their subgradient sets (2.2) and (2.3) to be polyhedral; see [45, Section 10E].

We start with a local reduction lemma that describes a general setting where the second-order chain rule (3.18) holds as equality. Then we show that it is the case for some major classes of fully amenable functions under the second-order qualification condition (3.19).

**Lemma 4.1 (local reduction to full rank condition).** *Let  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be a strongly amenable composition at  $\bar{x}$  represented as  $\varphi = \theta \circ h$  near  $\bar{x}$  with  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$  and  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ . Denote by  $S(z)$  the subspace parallel to the affine hull  $\text{aff } \partial\theta(z) \subset \mathbb{R}^m$ , and let  $L$  be any subspace of  $\mathbb{R}^m$  satisfying the inclusion*

$$(4.3) \quad L \supset S(z) \text{ for all } z \in \mathbb{R}^m \text{ sufficiently close to } \bar{z} = h(\bar{x}).$$

*Then we have the exact second-order chain rule formula (4.2).*

**Proof.** Let  $\dim L = s \leq m$ . It is easy to see that  $s = m$  corresponds to the full rank condition on  $\nabla h(\bar{z})$ , and thus the second-order chain rule (4.2) follows from Theorem 3.1. Suppose now that  $s < m$ , and let  $A$  be the matrix of a linear isometry from  $\mathbb{R}^m$  into  $\mathbb{R}^s \times \mathbb{R}^{m-s}$  under which  $AL = \mathbb{R}^s \times \{0\}$ . Denoting  $P := Ah$  and  $\vartheta := \theta A^{-1}$  gives us the representation  $\varphi = \vartheta \circ P$ . Hence the initial framework of the lemma can be reduced to one in which we have, in terms of  $P(x) = (p_1(x), \dots, p_m(x))$  and  $z = Ax$ , the implication

$$\left. \begin{array}{l} z \text{ sufficiently close to } \bar{z} \\ v = (v_1, \dots, v_m) \in \partial\vartheta(z) \end{array} \right\} \implies v_{s+1} = 0, \dots, v_m = 0.$$

This means that in analyzing  $\partial\varphi$  locally via  $\vartheta$  and  $P$  it is possible to pass with no loss of generality to the “submapping”

$$P_0: x \mapsto (p_1(x), \dots, p_s(x)),$$

since only  $p_1, \dots, p_s$  are active locally while  $p_{s+1}, \dots, p_m$  do not matter in the implication

$$y \in \partial\varphi(x) \implies \exists v \in \partial\vartheta(P(x)) \text{ such that } \nabla P(x)^* v = y.$$

It suffices therefore to impose a full rank ( $= s$ ) condition on  $P_0$  and invoke the second-order chain rule in equality form (4.2) from Theorem 3.1. The result of the lemma can be then translated back to the original context of  $\varphi = \theta \circ h$  through  $A^{-1}$ .  $\triangle$

Observe that the full rank condition on  $P_0$  comes out in the context of Lemma 4.1 from

$$L \cap \ker \nabla h(x)^* = \{0\}.$$

We show in what follows that a subspace  $L$  in (4.3) is provided by the second-order qualification condition (3.19) when  $\varphi = \theta \circ h$  belongs to major classes of fully amenable compositions.

To proceed in this direction, let us first calculate the second-order subdifferential  $\partial^2\theta(\bar{z}, \bar{v})(0)$  from the left-hand side of (3.19) for general convex piecewise linear-quadratic functions  $\theta$ , which is certainly of its own interest.

**Theorem 4.2 (calculating the second-order subdifferential of convex linear-quadratic functions).** *Let  $\varphi = \theta \circ h$  be a fully amenable composition at  $\bar{x}$ , let  $M: \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  be a set-valued mapping defined in (3.25), and let  $S(z)$  be a subspace of  $\mathbb{R}^m$  parallel to the affine hull*

aff  $\partial\theta(z)$  for  $z$  near  $\bar{z} = h(\bar{x})$ . Then for any sufficiently small neighborhood  $O$  of  $\bar{z}$  there is a finite union of the subspaces  $S(z)$  such that

$$(4.4) \quad \partial^2\theta(\bar{z}, \bar{v})(0) = \bigcup_{z \in O} S(z) \quad \text{whenever } \bar{v} \in M(\bar{x}, \bar{y}).$$

**Proof.** As mentioned in the proof of Theorem 3.3, the mapping  $M$  from (3.25) is closed-graph around  $(\bar{x}, \bar{y}, \bar{z})$  and uniformly bounded around  $(\bar{x}, \bar{y})$  for strongly amenable compositions. Since  $\theta \circ h$  is fully amenable, the set  $M(x, y)$  is also polyhedral for all  $(x, y)$  sufficiently close to  $(\bar{x}, \bar{y})$ .

Fix any  $\bar{v} \in M(\bar{x}, \bar{y})$ . Since  $\theta$  is piecewise linear-quadratic, its graph  $G := \text{gph } \theta$  is *piecewise polyhedral*, i.e., it is the union of finitely many polyhedral sets in  $\mathbb{R}^m$ . Using this and taking formulas (2.5) and (2.6) into account, we find a neighborhood  $W$  of  $(\bar{z}, \bar{v})$  such that

$$(4.5) \quad N_G(\bar{z}, \bar{v}) = \bigcup \left\{ \widehat{N}_G(z, v) \mid (z, v) \in G \cap W \right\} = \bigcup \left\{ T_G(z, v)^* \mid (z, v) \in G \cap W \right\},$$

where only *finitely many* cones (all of them are polyhedral) occur in the unions. Therefore

$$(4.6) \quad w \in \partial^2\theta(\bar{z}, \bar{v})(0) \iff \exists (z, v) \in G \cap W \quad \text{with } (w, 0) \in T_G(z, v)^*.$$

On the other hand, we have  $T_G(z, v) = \text{gph } (D\partial\theta)(z, v)$  by definition of the *graphical derivative*  $D$  of a set-valued mapping, and furthermore

$$(D\partial\theta)(z, v) = \partial\left(\frac{1}{2}d^2\theta(z, v)\right)$$

via the *second subderivative* of the function  $\theta$  under consideration; see [45, Theorem 13.40 and Proposition 13.32] for more details. Hence it ensures that

$$\text{dom } (D\partial\theta)(z, v) = \text{dom } d^2\theta(z, v) = N_{\partial\theta(z)}(v)$$

by [45, Theorem 13.14]. Employing now (4.5) and (4.6) gives us the representations

$$(4.7) \quad \begin{aligned} \partial^2\theta(\bar{z}, \bar{v})(0) &= \bigcup_{(z, v) \in G \cap W} \left[ \text{dom } (D\partial\theta)(z, v) \right]^* = \bigcup_{(z, v) \in G \cap W} \left[ N_{\partial\theta(z)}(v) \right]^* \\ &= \bigcup_{(z, v) \in G \cap W} T_{\partial\theta(z)}(v), \end{aligned}$$

where only finitely many sets are taken in the unions. Pick  $v \in \partial\theta(z)$  and find, by the polyhedrality of the subgradient sets  $\partial\theta(z)$  and the construction of the subspaces  $S(z)$ , a vector  $v' \in \text{ri } \partial\phi(z)$  arbitrary close to  $v$  and get for all such  $v'$  the relationships

$$T_{\partial\theta(z)}(v') = S(z) \supset T_{\partial\theta(z)}(v),$$

which imply by (4.7) the equality

$$\partial^2\theta(\bar{z}, \bar{v})(0) = \bigcup S(z),$$

where the finite union of *subspaces* are taken over  $z$  such that  $(z, v) \in G \cap W$  for some  $v$ .

So far we focused our analysis on a particular point  $\bar{v} \in M(\bar{x}, \bar{y})$  and an associated neighborhood  $W$  of  $(\bar{z}, \bar{v})$ . Since the mapping  $M$  is of closed graph and locally bounded, the set  $M(\bar{x}, \bar{y})$  can be covered by *finitely many* of such neighborhoods. This allows us to obtain (4.4) and complete the proof of the theorem.  $\triangle$

Note that representation (4.4) held for outer functions of any fully amenable composition is not generally sufficient for applying Lemma 4.1 and deducing thus the second-order chain rule as

equality (4.2) from the second-order qualification condition (3.19). To proceed in this direction, we need to get just *one subspace* in representation (4.4), which serves all  $\bar{v} \in M(\bar{x}, \bar{y})$  therein. The next result shows that it can be done in the case of *piecewise linear* outer functions in fully amenable compositions  $\varphi = \theta \circ h$ .

**Theorem 4.3 (second-order chain rule for fully amenable compositions with piecewise linear outer functions).** *Let  $\varphi = \theta \circ h$  be a fully amenable composition at  $\bar{x}$ , where  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  is (convex) piecewise linear. Assume that the second-order qualification condition (3.19) is satisfied. Then we have the exact second-order chain rule formula (4.2).*

**Proof.** It is not hard to check that for convex piecewise linear functions  $\varphi$  we have the inclusion  $\partial\theta(z) \subset \partial\theta(\bar{z})$  for any neighborhood  $O$  of  $\bar{z}$  sufficiently small. This implies that  $S(z) \subset S(\bar{z})$  whenever  $z \in O$  and hence ensures the equality

$$(4.8) \quad \partial^2\theta(\bar{z}, \bar{v})(0) = S(\bar{z})$$

by formula (4.4) from Theorem 4.2. Employing the latter subdifferential representation in the second-order qualification condition (3.19) gives us

$$(4.9) \quad S(z) \cap \ker \nabla h(\bar{x})^* = \{0\} \quad \text{for all } z \in O.$$

Recall that the subspace  $S(z)$  consists of all vectors  $\lambda(v' - v)$  such that  $\lambda \in \mathbb{R}$  and  $\lambda, \lambda' \in \partial\theta(z)$ . Hence the second-order qualification condition (4.9) is equivalent to the following: there exist neighborhoods  $U$  of  $\bar{z}$  and  $V$  of  $\bar{y}$  such that

$$\left[ x \in U, y \in V, v, v' \in M(x, y) \right] \implies v = v'.$$

On other words, the latter means that the mapping  $M$  from (3.25) is *single-valued* on the subdifferential graph  $\text{gph } \partial\varphi$  around  $(\bar{x}, \bar{y})$  in the case under consideration; thus it is continuous as well. The result of the theorem follows now from Lemma 4.1.  $\triangle$

Next we consider a major subclass of piecewise linear-quadratic outer functions in fully amenable compositions given by

$$(4.10) \quad \theta(z) := \sup_{v \in C} \left\{ \langle v, z \rangle - \frac{1}{2} \langle v, Qv \rangle \right\},$$

where  $C \subset \mathbb{R}^m$  is a nonempty polyhedral set, and where  $Q \in \mathbb{R}^{m \times m}$  is a symmetric positive-semidefinite matrix. Functions of this class are useful in many aspects of variational analysis and optimization; in particular, as penalty expressions in composite formats of optimization; see, e.g., [45] and the references therein. By definition (4.10) we see that  $\theta$  in (4.10) is proper, convex, and piecewise linear-quadratic (piecewise linear when  $Q = 0$ ) with the conjugate representation

$$(4.11) \quad \theta(z) = (\delta_C + j_Q)^*(z) \quad \text{for } j_Q(v) := \frac{1}{2} \langle v, Qv \rangle.$$

In the following theorem we calculate the second-order subdifferential of functions  $\theta$  from (4.10) and justifies the equality second-order chain rule formula (4.2) for fully amenable compositions  $\varphi = \theta \circ h$  with outer functions of this type.

**Theorem 4.4 (second-order calculus rule for a major subclass of fully amenable compositions).** *Let  $\varphi = \theta \circ h$  be a fully amenable composition at  $\bar{x}$  with  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  of class (4.10) satisfying the assumptions made above. Suppose also that the second-order qualification condition (3.19) is satisfied. Then we have the exact second-order chain rule formula (4.2).*

**Proof.** It follows from (4.11) that the conjugate to  $\theta$  is  $\theta^* = \delta_C + j_Q$ . Observe also that

$$(4.12) \quad w \in \partial^2\theta(\bar{z}, \bar{v})(u) \iff -u \in \partial^2\theta^*(\bar{v}, \bar{z})(-w).$$

Furthermore, we have by the calculations in [45, Example 11.18] that

$$(4.13) \quad \partial\theta^*(v) = N_C(v) + Qv, \quad v \in C,$$

and hence  $\bar{z} \in \partial\theta^*(\bar{v}) \iff \bar{z} - Q\bar{v} \in N_C(\bar{v})$ . Using this and definition (2.11) of the second-order subdifferential and then applying the coderivative sum rule [29, Theorem 1.62] to (4.13) give us

$$\partial^2\theta^*(\bar{v}, \bar{u})(-w) + D^*N_C(\bar{v}, \bar{z} - Q\bar{v})(-w) + \nabla j_Q(-w).$$

Since  $\nabla j_Q(-w) = Qw$ , the latter implies in turn that

$$(4.14) \quad -u \in \partial^2\theta^*(\bar{v}, \bar{z})(-w) \iff Qw - u \in \partial^2\delta_C(\bar{v}, \bar{z} - Q\bar{v})(-w).$$

Employing (4.12), (4.14) and proceeding similarly to the consideration in Example 3.5 above, we get from (4.14) the *exact formula for calculating the second-order subdifferential*:

$$(4.15) \quad u \in \partial^2\theta(\bar{z}, \bar{v})(w) \iff \begin{cases} \exists \text{ closed faces } K_1 \supset K_2 \text{ of } K \\ \text{with } w \in K_1 - K_2, Qw - u \in (K_1 - K_2)^*, \end{cases}$$

where  $K$  is the critical cone for  $C$  at  $\bar{v}$  with respect to  $\bar{z} - Q\bar{v}$  given by

$$K = T_C(\bar{v}) \cap (\bar{z} - Q\bar{v})^\perp.$$

The positive-semidefiniteness of the matrix  $Q$  yields that

$$0 \geq \langle w, Qw \rangle \iff Qw = 0 \iff w \in \ker Q,$$

which allows us to deduce from (4.15) that

$$\begin{aligned} w \in \partial^2\theta(\bar{z}, \bar{v})(0) &\iff \exists K_1 \supset K_2 \text{ with } w \in K_1 - K_2, Qw \in (K_1 - K_2)^* \\ &\iff \exists K_1 \supset K_2 \text{ with } w \in (\ker Q) \cap (K_1 - K_2) \\ &\iff w \in (\ker Q) \cap (K - K). \end{aligned}$$

Thus the set  $\partial^2\theta(\bar{z}, \bar{v})(0)$  is a *subspace* in  $\mathbb{R}^m$ . Substituting it finally into the second-order subdifferential condition (3.19) and taking into account that  $\partial\theta(z) = (N_C + Q)^{-1}(z)$  by [45, Example 11.18], we arrive at the second-order equality chain rule (4.2) similarly to the proof of Theorem 4.3 based on the application of Lemma 4.1.  $\triangle$

## 5 Applications to Tilt Stability in Nonlinear and Extended Nonlinear Programming

The second-order chain rules and subdifferential calculations obtained in Sections 3 and 4 are undoubtedly useful in any settings where the second-order subdifferential (2.11) and its partial counterparts are involved; see the discussions and references in Section 1. In this section we confine ourselves to the usage of second-order chain rules for deriving *full characterizations* of *tilt-stable local minimizers* in some important classes of constrained optimization problems. It requires applying *equality-type* formulas of the second-order subdifferential calculus.

The notion of tilt-stable minimizers was introduced by Poliquin and Rockafellar [39] in order to characterize strong manifestations of optimality that support computational work via

the study of how local optimal solutions react to shifts (tilt perturbations) of the data. Recall that a point  $\bar{x}$  is a *tilt-stable local minimizer* of the function  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  finite at  $\bar{x}$  if there is  $\gamma > 0$  such that mapping

$$M: y \mapsto \operatorname{argmin} \left\{ \varphi(x) - \varphi(\bar{x}) - \langle y, x - \bar{x} \rangle \mid \|x - \bar{x}\| \leq \gamma \right\}$$

is single-valued and Lipschitz continuous on some neighborhood of  $y = 0$  with  $M(0) = \bar{x}$ .

It is proved in [39, Theorem 1.3] that for  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  having  $0 \in \partial\varphi(\bar{x})$  and such that  $\varphi$  is both prox-regular and subdifferentially continuous at  $\bar{x}$  for  $\bar{y} = 0$ , the point  $\bar{x}$  is a tilt-stable local minimizer of  $\varphi$  if and only if the second-order subdifferential mapping  $\partial^2\varphi(\bar{x}, 0): \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is *positive-definite* in the sense that

$$(5.1) \quad \langle w, u \rangle > 0 \quad \text{whenever} \quad w \in \partial^2\varphi(\bar{x}, 0)(u) \quad \text{with} \quad u \neq 0.$$

The aforementioned properties of prox-regularity and subdifferential continuity introduced in [38] (see also [45, Definitions 13.27 and 13.28]) hold for broad classes of “nice” functions encountered in variational analysis and optimization. In particular, both properties are satisfied, at all points of a neighborhood of  $\bar{x}$  for any function strongly amenable at  $\bar{x}$ ; see [45, Proposition 13.32].

Our subsequent goal is to extend the characterization of tilt-stable local minimizers from [39] to favorable classes of *constrained optimization* problems. To proceed, we use the following *composite format* of optimization known as *extended nonlinear programming* (ENLP); see [44, 45]:

$$(5.2) \quad \text{minimize} \quad \varphi(x) := \varphi_0(x) + \theta(\varphi_1(x), \dots, \varphi_m(x)) = \varphi_0(x) + (\theta \circ \Phi)(x) \quad \text{over} \quad x \in \mathbb{R}^n,$$

where  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  is an extended-real-valued function, and where  $\Phi(x) := (\varphi_1(x), \dots, \varphi_m(x))$  is a mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ . Written in the unconstrained format, problem (5.2) is actually a problem of constrained optimization with the set of feasible solutions given by

$$X := \{x \in \mathbb{R}^n \mid (\varphi_1(x), \dots, \varphi_m(x)) \in Z\} \quad \text{for} \quad Z := \{z \in \mathbb{R}^m \mid \theta(z) < \infty\}.$$

In other words, problem (5.2) can be equivalently represented in the form

$$(5.3) \quad \text{minimize} \quad \varphi_0(x) + \delta_Z(\Phi(x)) \quad \text{over} \quad x \in \mathbb{R}^n \quad \text{with} \quad Z = \operatorname{dom} \theta$$

via the indicator function of the feasible set. As argued in [44], the composite format (5.2), or (5.3), is a convenient general framework from both theoretical and computational viewpoints to accommodate a variety of particular models in constrained optimization. Note that the conventional problem of nonlinear programming with  $s$  inequality constraints and  $m - s$  equality constraints can be written in form (5.3), where  $Z = \mathbb{R}_-^s \times \{0\}^{m-s}$ .

Our first result provides a complete second-order characterization of tilt-stable minimizers  $\bar{x}$  for a general class of problems (5.2) under full rank of the Jacobian matrix  $\nabla\Phi(\bar{x})$ .

**Theorem 5.1 (characterization of tilt-stable minimizers for constrained problems with full rank condition).** *Let  $\bar{x} \in X$  be a feasible solution to (5.2) such that  $\varphi_0$  and  $\Phi$  are smooth around  $\bar{x}$  with their derivatives strictly differentiable at  $\bar{x}$ , that  $\operatorname{rank} \nabla\Phi(\bar{x}) = m$ , and that  $\theta$  is prox-regular and subdifferentially continuous at  $\bar{z} := \Phi(\bar{x})$  for the (unique) vector  $\bar{v} \in \mathbb{R}^m$  satisfying the relationships*

$$(5.4) \quad \bar{v} \in \partial\theta(\bar{z}) \quad \text{and} \quad \nabla\Phi(\bar{x})^* \bar{v} = -\nabla\varphi_0(\bar{x}).$$

*Then  $\bar{x}$  with  $-\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^* \partial\theta(\bar{z})$  is a tilt-stable local minimizer of (5.2) if and only if the mapping  $T: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  given by*

$$(5.5) \quad T(u) := \nabla^2 \langle \bar{v}, \Phi \rangle(\bar{x})(u) + \nabla\Phi(\bar{x})^* \partial^2\theta(\bar{z}, \bar{v})(\nabla\Phi(\bar{x})u), \quad u \in \mathbb{R}^n,$$

*is positive-definite in the sense of (5.1).*

**Proof.** Since  $\varphi_0$  and  $\Phi$  are smooth around  $\bar{x}$  and  $\nabla\Phi(\bar{x})$  has full rank  $m$ , it follows from the first-order subdifferential sum and chain rules of [45, Corollary 10.9 and Exercise 10.7] that

$$0 \in \partial\varphi(\bar{x}) \iff -\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^* \partial\theta(\bar{z})$$

for  $\varphi$  in (5.2). Furthermore, these rules and the definitions of prox-regularity and subdifferential continuity in [45] imply that the latter properties of  $\varphi$  at  $\bar{x}$  for  $0 \in \partial\varphi(\bar{x})$  are equivalent to the corresponding properties of  $\theta$  at  $\bar{z}$  for  $\bar{v}$  satisfying (5.4).

It remains to check therefore that the positive-definiteness (5.1) of  $\partial^2\varphi(\bar{x}, 0)$  is equivalent to that of  $T$  in (5.5). We show in fact that  $\partial^2\varphi(\bar{x}, 0)(u) = T(u)$  for all  $u \in \mathbb{R}^n$ . Indeed, using the second-order chain rule from [29, Proposition 1.121] in (5.2) gives us

$$(5.6) \quad \partial^2\varphi(\bar{x}, 0)(u) = \nabla^2\varphi_0(\bar{x})u + \partial^2(\theta \circ \Phi)(\bar{x}, -\nabla\varphi_0(\bar{x}))(u), \quad u \in \mathbb{R}^n.$$

To complete the proof of the theorem, we finally apply the exact second-order chain rule from Theorem 3.1 to the composition  $\theta \circ \Phi$  in the latter equality.  $\triangle$

Next we address the conventional model of *nonlinear programming* (NLP) with smooth data:

$$(5.7) \quad \text{minimize } \varphi_0(x) \text{ subject to } \varphi_i(x) = \begin{cases} \leq 0 & \text{for } i = 1, \dots, s, \\ = 0 & \text{for } i = s+1, \dots, m. \end{cases}$$

As mentioned above, problem (5.7) can be written in form (5.3) with  $Z = \mathbb{R}_-^s \times \{0\}^{m-s}$ . For this problem, the full rank condition of Theorem 5.1 corresponds to: the gradients

$$(5.8) \quad \nabla\varphi_1(\bar{x}), \dots, \nabla\varphi_m(\bar{x}) \text{ are linearly independent.}$$

Actually, since our analysis is local, we can drop in what follows any inactive inequality constraints from the picture and thus reduce with no loss of generality to having

$$(5.9) \quad \varphi_i(\bar{x}) = 0 \text{ for all } i = 1, \dots, m.$$

The full rank condition (5.8) in case (5.9) is then the classical (LICQ): *the active constraint gradients at  $\bar{x}$  are linearly independent*.

To proceed further, consider the Lagrangian function in (5.7) given by

$$L(x, \lambda) := \varphi_0(x) + \sum_{i=1}^m \lambda_i \varphi_i(x) \text{ with } \lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$$

and remember that, for any local optimal solution  $\bar{x}$  to (5.7), the LICQ at  $\bar{x}$  ensures the existence of a *unique* multiplier vector  $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_m) \in \mathbb{R}_+^s \times \mathbb{R}^{m-s}$  such that

$$(5.10) \quad \nabla_x L(\bar{x}, \bar{\lambda}) = \nabla\varphi_0(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla\varphi_i(\bar{x}) = 0.$$

Recall that the *strong second-order optimality condition* (SSOC) holds at  $\bar{x}$  if

$$(5.11) \quad \langle u, \nabla_{xx}^2 L(\bar{x}, \bar{\lambda})u \rangle > 0 \text{ for all } 0 \neq u \in S,$$

where the subspace  $S \subset \mathbb{R}^n$  is given by

$$S := \{u \in \mathbb{R}^n \mid \langle \nabla\varphi_i(\bar{x}), u \rangle = 0 \text{ whenever } i = 1, \dots, m\}.$$

Note that (5.11) is also known as the “strong second-order sufficient condition” for local optimality. The following theorem shows that, in the setting under consideration, the SSOC is *necessary and sufficient* for the tilt stability of local minimizers.

**Theorem 5.2 (characterization of tilt-stable local minimizers for NLP).** *Let  $\bar{x}$  be a feasible solution to (5.7) such that all the functions  $\varphi_i$  for  $i = 0, \dots, m$  are smooth around  $\bar{x}$  with their derivatives strictly differentiable at  $\bar{x}$  and that the LICQ is satisfied at this point. Then we have the following assertions:*

- (i) *If  $\bar{x}$  is a tilt-stable local minimizer of (5.7), then SSOC (5.11) holds at  $\bar{x}$  with the unique multiplier vector  $\bar{\lambda} \in \mathbb{R}_+^s \times \mathbb{R}^{m-s}$  satisfying (5.10).*
- (ii) *Conversely, the validity of SSOC at  $\bar{x}$  with  $\bar{\lambda} \in \mathbb{R}_+^s \times \mathbb{R}^{m-s}$  satisfying (5.10) implies that  $\bar{x}$  is a tilt-stable local minimizer of (5.7).*

**Proof.** As mentioned above, the LICQ corresponds to the full rank condition of Theorem 5.1 in the setting under consideration. The prox-regularity and subdifferential continuity of the indicator function  $\theta = \delta_Z$  with  $Z = \mathbb{R}_-^s \times \{0\}^{m-s}$  follow from its convexity [45, Example 13.30]. Let us next represent the mapping  $T$  in (5.5) via the initial data of problem (5.7). It is easy to see that  $T(u)$  reduces in this case to

$$T(u) = \nabla_{xx}^2 L(\bar{x}, \bar{\lambda})u + \nabla \Phi(\bar{x})^* \partial^2 \delta_Z(0, \bar{\lambda})(\nabla \Phi(\bar{x})u),$$

with  $\Phi = (\varphi_1, \dots, \varphi_m)$  and  $Z = \mathbb{R}_-^s \times \{0\}^{m-s}$ , provided that the first-order condition (5.10) is satisfied, which is of course the case when  $\bar{x}$  is a local minimizer of (5.7). Thus the positive-definiteness of  $T(u)$  amounts to

$$(5.12) \quad u \neq 0, w \in \partial^2 \delta_Z(0, \bar{\lambda})(u) \implies \langle u, \nabla_{xx}^2 L(\bar{x}, \bar{\lambda})u \rangle + \langle w, \nabla \Phi(\bar{x})u \rangle > 0.$$

To proceed, we calculate the second-order subdifferential  $\partial^2 \delta_Z(0, \bar{\lambda})$  in (5.12) by using formula (3.39) presented and discussed in Example 3.5. Observe that the critical cone in this situation is  $K = Z \cap \bar{\lambda}^\perp$ . It follows directly from (3.39) that

$$w \in \partial^2 \delta_Z(0, \bar{\lambda})(u) \iff \begin{cases} \text{there exist closed faces } K_1 \subset K_2 \text{ of } K \\ \text{with } -\nabla \Phi(\bar{x})u \in K_1 - K_2, w \in (K_2 - K_1)^*. \end{cases}$$

The latter implies in turn that

$$\min_{w \in \partial^2 \delta_Z(0, \bar{\lambda})(u)} \langle w, \nabla \Phi(\bar{x})u \rangle = 0 \quad \text{for all } u \in \text{dom } \partial^2 \delta_Z(0, \bar{\lambda})$$

with the subdifferential domain representation

$$\text{dom } \partial^2 \delta_Z(0, \bar{\lambda}) = \bigcup \left\{ (K_1 - K_2) \mid K_1 \subset K_2 \text{ closed faces of } K \right\} = K - K.$$

Substituting this into (5.12) and taking into account the forms of the critical cone  $K$  as well as the subspace  $S$  in (5.11), we conclude that the positive-definiteness condition (5.12) is equivalent to the strong second-order optimality condition (5.11) provided that (5.10) holds at  $\bar{x}$ .

Having this in hand and using Theorem 5.1 allow us to justify both assertions in (i) and (ii). Indeed, since every tilt-stable local minimizers  $\bar{x}$  is a standard local minimizer, it satisfies the first-order necessary optimality condition (5.10) under the assumed LICQ at  $\bar{x}$ . Thus the SSOC holds at this point, which is the assertion in (i). The validity of the converse assertion (ii) follows from the equivalence between (5.12) and (5.11) under (5.10) proved above.  $\triangle$

The obtained characterization of tilt-stable minimizers for NLP leads us to comparing this notion with the classical Robinson's notion of *strong regularity* [41] of parameterized variational inequalities associated with the KKT conditions for NLP (5.7). Complete characterizations of strong regularity for NLP are derived in [9]; see also the references therein.

**Corollary 5.3 (comparing tilt-stability and strong regularity).** *Under the assumptions of Theorem 5.2, the tilt-stability of local minimizers for (5.7) is equivalent to the strong regularity of the variational inequality associated with the KKT conditions for (5.7).*

**Proof.** It follows directly from Theorem 5.2 and the characterization of strong regularity obtained in [9, Theorem 5 and Theorem 6].  $\triangle$

It is not hard to check that the strong regularity of the KKT system directly implies the LICQ at the corresponding solution of (5.7). On the other hand, the LICQ requirement arising from the full rank condition of Theorem 5.1 is essential for the SSOC characterization of tilt-stability in Theorem 5.2. Furthermore, even imposing the seemingly less restrictive second-order qualification condition (3.19) needed for deriving the second-order chain rule *unavoidably* leads us to the LICQ requirement for NLP, since the latter class is represented via fully amenable compositions with piecewise linear outer functions  $\theta$  in the composite format (5.2). This follows from the results of Section 4 and is reflected in the next theorem.

**Theorem 5.4 (characterizing tilt-stable minimizers for constrained problems described by fully amenable compositions).** *Let  $\bar{x}$  be a feasible solution to (5.2) such that  $\varphi_0$  smooth around  $\bar{x}$  with the strictly differentiable derivative at  $\bar{x}$  and that the composition  $\theta \circ \Phi$  is fully amenable at  $\bar{x}$  with the outer function  $\theta: \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  of the following types:*

- *either  $\theta$  is piecewise linear,*
- *or  $\theta$  is of class (4.10) with a nonempty polyhedral set  $C \subset \mathbb{R}^m$  and a symmetric positive-semidefinite matrix  $Q \in \mathbb{R}^m \times \mathbb{R}^m$ .*

*Assume further that the second-order qualification condition (3.19) holds at  $\bar{x}$ , where  $v = -\bar{v}$  is the unique vector satisfying (5.4) with  $\bar{z} = \Phi(\bar{x})$ . Then  $\bar{x}$  with  $-\nabla\varphi_0(\bar{x}) \in \nabla\Phi(\bar{x})^*\partial\theta(\bar{z})$  is a tilt-stable local minimizer of (5.2) if and only if the mapping  $T: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  defined in (5.5) is positive-definite in the sense of (5.1), where the second-order subdifferential  $\partial^2\theta(\bar{z}, \bar{v})$  is calculated by formulas (4.8) and (4.15), respectively.*

**Proof.** Observe first that in both cases under consideration the composition  $\theta \circ \Phi$  is prox-regular and subdifferentially continuous at any point  $x$  around  $\bar{x}$  by [45, Proposition 13.32]; hence the same holds for the function  $\varphi$  from (5.2). It follows from Theorems 4.3 and 4.4 that, under the validity of the second-order qualification condition (3.19), we have the unique vector  $\bar{v}$  satisfying (5.4) and the second-order chain rule

$$(5.13) \quad \partial^2(\theta \circ \Phi)(\bar{x}, -\nabla\varphi_0(\bar{x}))(u) = \left( \nabla^2\langle \bar{v}, \Phi \rangle(\bar{x})u + \nabla\Phi(\bar{x})^*\partial^2\theta(\bar{z}, \bar{v})(\nabla\Phi(\bar{x})u) \right)$$

for all  $u \in \mathbb{R}^n$  when  $\theta$  belongs to one of the classes considered in this theorem. Substituting further (5.13) into formula (5.6) due to the second-order sum rule from [29, Proposition 1.121] allows us to justify that

$$\partial^2\varphi(\bar{x}, 0) = T(u) \text{ whenever } u \in \mathbb{R}^n,$$

and thus the positive-definiteness of the mapping  $T$  from (5.5) fully characterizes the tilt-stability of the local minimizer  $\bar{x}$  of (3.1) in both cases of  $\theta$  under consideration with the formulas for calculating of  $\partial^2\theta(\bar{z}, \bar{v})$  derived in the proofs of Theorems 4.3 and 4.4, respectively.  $\triangle$

We conclude the paper with the following two final remarks.

**Remark 5.5 (sufficient conditions for tilt-stable local minimizers).** The second-order chain rule (3.18) of the *inclusion* type derived in Theorem 3.3 for *strongly amenable* compositions



and the second-order sum rule inclusions obtained in [29, Theorem 3.73] allow us to establish general *sufficient* conditions for tilt-stable local minimizers in large classes of constrained optimization problems written in the composite format (5.2). Indeed if, in addition to the hypotheses of Theorem 3.3 for the composition  $\theta \circ \Phi$  in (5.2), we assume that the function  $\varphi_0$  is, e.g., of class  $\mathcal{C}^{1,1}$  around  $\bar{x}$ , then we have by the second-order sum rule from [29, Theorem 3.73(i)] and the chain rule of Theorem 3.3 the fulfillment of the inclusion

$$(5.14) \quad \partial^2 \varphi(\bar{x}, 0)(u) \subset T(u), \quad u \in \mathbb{R}^n,$$

for  $\varphi$  from (5.2) and  $T$  from (5.5). The prox-regularity and subdifferential continuity of such functions  $\varphi$  follow, under the assumptions made, from [45, Proposition 13.32 and Proposition 13.34] and first-order subdifferential calculus rules. Thus inclusion (5.14) ensures that the positive-definiteness of  $T$  implies the one of  $\partial^2 \varphi(\bar{x}, 0)$ , and the former is therefore a sufficient condition for the tilt-stability of local minimizers of (5.2).

**Remark 5.6 (full stability of local minimizers).** Developing the concept of tilt stability, Levy, Poliquin and Rockafellar [19] introduced the notion of *fully stable* local minimizers of general optimization problems of the type

$$(5.15) \quad \text{minimize } \varphi(x, u) - \langle v, x \rangle \quad \text{over } x \in \mathbb{R}^n$$

with respect to both “basic” perturbations  $u$  and “tilt” perturbations  $v$ . The main result of that paper [19, Theorem 2.3] establishes a *complete characterization* of fully stable local minimizers of (5.15) via the positive-definiteness of the *extended partial second-order subdifferential* (2.13) of  $\varphi$ . Similarly to the results of this section for tilt-stable minimizers of constraint optimization problems written in the composite format (5.2), we can derive characterizations as well as sufficient conditions for fully stable minimizers of (5.15) based on [19, Theorem 2.3] and the second-order chain rules for the partial second-order counterpart (2.13) obtained above. Our ongoing research project is to comprehensively elaborate these developments on full stability in constrained optimization and its applications.

**Remark 5.7 (tilt stability and partial smoothness).** After completing this paper, we became aware of the concurrent work by Lewis and Zhang [21] related to second-order subdifferentials (generalized Hessians) and tilt stability. The main results of [21] provide calculations of the basic second-order construction from Definition 2.1 for  $\mathcal{C}^2$ -partly smooth functions on  $\mathcal{C}^2$ -smooth manifolds and then characterize tilt stability in such settings via strong criticality and local quadratic growth.

## References

- [1] L. BAN, B. S. MORDUKHOVICH and W. SONG, *Lipschitzian stability of parametric variational inequalities over generalized polyhedra in Banach spaces*, Nonlinear Anal., 74 (2011), pp. 441–461.
- [2] J. F. BONNANS and A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Series in Operations Research, Springer, New York, 2000.
- [3] J. M. BORWEIN and Q. J. ZHU, *Techniques of Variational Analysis*, CMS Books in Mathematics, Vol. 20, Springer, New York, 2005.
- [4] N. H. CHIEU, T. D. CHUONG, J.-C. YAO and N. D. YEN, *Characterizing convexity of a function by its Fréchet and limiting second-order subdifferentials*, Set-Valued Var. Anal., 19 (2011), pp. 75–96.
- [5] N. H. CHIEU and N. D. HUY, *Second-order subdifferentials and convexity of real-valued functions*, Nonlinear Anal., 74 (2011), pp. 154–160.

- [6] N. H. CHIEU and N. T. Q. TRANG, *Coderivatives and monotonicity of continuous mappings*, Taiwan. J. Math., to appear (2011).
- [7] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, CMS Books in Mathematics, Vol. 1, Wiley, New York, 1983.
- [8] G. COLOMBO, R. HENRION, N. D. HOANG and B. S. MORDUKHOVICH, *Optimal control of the sweeping process*, Dynam. Cont. Disc. Impul. Syst., Ser. B, 19 (2012).
- [9] A. L. DONTCHEV and R. T. ROCKAFELLAR, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.
- [10] A. L. DONTCHEV and R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer Monographs in Mathematics, Springer, Dordrecht, 2009.
- [11] A. C. EBERHARD and R. WENCZEL, *A study of tilt-stable optimality and sufficient conditions*, Nonlinear Anal., 75 (2012), No. 3.
- [12] R. HENRION, B. S. MORDUKHOVICH and N. M. NAM, *Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities*, SIAM J. Optim., 20 (2010), pp. 2199–2227.
- [13] R. HENRION, J. V. OUTRATA and T. SUROWIEC, *On the coderivative of normal cone mappings to inequality systems*, Nonlinear Anal., 71 (2009), pp. 1213–1226.
- [14] R. HENRION, J. V. OUTRATA and T. SUROWIEC, *Analysis of M-stationarity points to an EPEC modeling oligopolistic competition in an electricity spot market*, ESAIM: Contr. Optim. Calc. Var., to appear (2012).
- [15] R. HENRION and W. RÖMISCH, *On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling*, Appl. Math., 52 (2007), pp. 473–494.
- [16] N. Q. HUY and J.-C. YAO, *Exact formulae for coderivatives of normal cone mappings to perturbed polyhedral convex sets*, J. Optim. Theory Appl., to appear (2012).
- [17] V. JEYAKUMAR and D. T. LUC, *Nonsmooth Vector Functions and Continuous Optimization*, Optimization and Its Applications, Vol. 10, Springer, New York, 2008.
- [18] A. B. LEVY and B. S. MORDUKHOVICH, *Coderivatives in parametric optimization*, Math. Program., 99 (2004), pp. 311–327.
- [19] A. B. LEVY, R. A. POLIQUIN and R. T. ROCKAFELLAR, *Stability of locally optimal solutions*, SIAM J. Optim., 10 (2000), pp. 580–604.
- [20] A. B. LEVY and R. T. ROCKAFELLAR, *Variational conditions and the proto-differentiation of partial subgradient mappings*, Nonlinear Anal., 26 (1996), pp. 1951–1964.
- [21] A. S. LEWIS and S. ZHANG, *Partial smoothness, tilt stability, and generalized Hessians*, preprint (2011).
- [22] B. S. MORDUKHOVICH, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech., 40 (1976), pp. 960–969.
- [23] B. S. MORDUKHOVICH, *Metric approximations and necessary optimality conditions for general classes of extremal problems*, Soviet Math. Dokl., 22 (1980), pp. 526–530.
- [24] B. S. MORDUKHOVICH, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design, D. A. Field and V. Komkov (eds.), Proceedings in Applied Mathematics, Vol. 58, SIAM, Philadelphia, 1992, pp. 32–46.
- [25] B. S. MORDUKHOVICH, *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 340 (1993), pp. 1–35.
- [26] B. S. MORDUKHOVICH, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl., 183 (1994), pp. 250–288.

- [27] B. S. MORDUKHOVICH, *Stability theory for parameteric generalized equations and variational inequalities via nonsmooth analysis*, Trans. Amer. Math. Soc., 343 (1994), pp. 609–658.
- [28] B. S. MORDUKHOVICH, *Calculus of second-order subdifferentials in infinite dimensions*, Control Cybernet., 31 (2002), pp. 557–573.
- [29] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330, Springer, Berlin, 2006.
- [30] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, II: Applications*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 331, Springer, Berlin, 2006.
- [31] B. S. MORDUKHOVICH and J. V. OUTRATA, *Second-order subdifferentials and their applications*, SIAM J. Optim., 12 (2001), pp. 139–169.
- [32] B. S. MORDUKHOVICH and J. V. OUTRATA, *Coderivative analysis of quasi-variational inequalities with applications to stability in optimization*, SIAM J. Optim., 18 (2007), pp. 389–412.
- [33] B. S. MORDUKHOVICH and Y. SHAO, *Nonsmooth sequential anaysis in Asplund spaces*, Trans. Amer. Math. Soc., 348 (1996), 1235–1280.
- [34] B. S. MORDUKHOVICH and B. WANG, *Restrictive metric regularity and generalizerd differential calculus in Banach spaces*, Int. J. Maths. Math. Sci., 50 (2004), pp. 2650–2683.
- [35] N. M. NAM, *Coderivatives of the normal cone mappings and Lipschitzian stability of parametric variational inequalities*, Nonlinear Anal., 73 (2010), pp. 2271–2282.
- [36] J. V. OUTRATA, *Mathematical programs with equilibrium constraints: Theory and numerical methods*, in Nonsmooth Mechanics of Solids, CISM Lecture Notes, Vol. 485, J. Haslinger and G. E. Stavroulakis (eds.), Springer, New York, 2006, pp. 221–274.
- [37] R. A. POLIQUIN and R. T. ROCKAFELLAR, *Amenable functions in optimization*, in Nonsmooth Optimization Methods and Applications, F. Giannessi (ed.), Gordon and Breach, Philadelphia, 1992, pp. 338–353.
- [38] R. A. POLIQUIN and R. T. ROCKAFELLAR, *Prox-regular functions in variational analysis*, Trans. Amer. Math. Soc., 348 (1996), pp. 1805–1838.
- [39] R. A. POLIQUIN and R. T. ROCKAFELLAR, *Tilt stability of a local minimum*, SIAM J. Optim., 8 (1998), pp. 287–299.
- [40] N. T. QUI, *Nonlinear perturbations of polyhedral normal cone mappings and affine variational inequalities*, J. Optim. Theory Appl., to appear (2012).
- [41] S. M. ROBINSON, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [42] S. M. ROBINSON, *Some continuity properties of polyhedral multifunctions*, Math. Programming Stud., 14 (1981), pp. 206–214.
- [43] R. T. ROCKAFELLAR, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. H. Poincaré: Analyse Non Linéaire, 2 (1985), 167–184.
- [44] R. T. ROCKAFELLAR, *Extended nonlinear programming*, in Nonlinear Optimization and Related Topics, G. Di Pillo and F. Giannessi (eds.), Applied Optimization, Vol. 36, Kluwer Academic Publishers, Dordrecht, 2000, pp. 381–399.
- [45] R. T. ROCKAFELLAR and R. J-B WETS, *Variational Analysis*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 317, Springer, Berlin, 2006.
- [46] W. SCHIROTZEK, *Nonsmooth Analysis*, Universitext, Springer, Berlin, 2007.
- [47] T. SUROWIEC, *Explicit Stationarity Conditions and Solution Characterization for Equilibrium Problems with Equilibrium Constraints*, Ph. D. Disseration, Humboldt University, Berlin, 2010.

- [48] J.-C. YAO and N. D. YEN, *Coderivative calculation related to a parametric affine variational inequality. Part 1: Basic calculation*, Acta Math. Vietnam., 34 (2009), pp. 157–172.
- [49] J.-C. YAO and N. D. YEN, *Coderivative calculation related to a parametric affine variational inequality. Part 2: Applications*, Pac. J. Optim., 5 (2009), pp. 493–506.
- [50] J. J. YE, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim., 10 (2000), pp. 943–962.
- [51] R. ZHANG, *Problems of hierarchical optimization in finite dimensions*, SIAM J. Optim., 4 (1994), pp. 521–536.